Investigating heterogeneity and mechanisms of resistance in paediatric B-cell non-Hodgkin lymphoma

Sorcha Forde

Homerton College

Department of Pathology

Supervised by **Suzanne D. Turner, Ph.D.**

This dissertation is submitted for the degree of Doctor of Philosophy

University of Cambridge, United Kingdom – September 2018

Summary

Investigating heterogeneity and mechanisms of resistance in paediatric B-cell non-Hodgkin lymphoma

Sorcha Forde

Paediatric mature B-cell non-Hodgkin lymphoma (B-NHL) is a highly aggressive disease with one of the fastest growth rates of any human malignancy. The subtypes are Burkitt and Burkitt-like lymphoma/leukaemia, diffuse large B-cell lymphoma and primary mediastinal B-cell lymphoma; with Burkitt lymphoma/leukaemia alone accounting for 38% of NHL diagnoses in patients under 15 years. Outcomes have improved significantly in recent years with the introduction of intensive multi-agent chemotherapy, followed by the addition of Rituximab, resulting in a survival rate exceeding 90%. Despite this, those patients who acquire resistance to chemotherapy and/or Rituximab have a dismal prognosis with few salvage therapy options. The aim of this thesis was to identify cell populations that are driving therapy resistance in order to develop targeted treatments for relapsed and therapy-refractory patients.

Utilising samples obtained from patients enrolled in the Inter-B-NHL ritux 2010 clinical trial, we established 5 paediatric Patient Derived Xenografts (PDXs) – to our knowledge the largest paediatric B-NHL resource of this type. These PDXs faithfully recapitulate tumour characteristics through passage and enable the identification of heterogeneous populations within these malignancies. One such population, known as the side population (SP), exhibits tumour initiating potential, expresses stemness-related genes and shows increased resistance to chemotherapy. *In vitro*, the SP produces soluble factors necessary for growth of the main population (MP) of cells (depleted of SP); indeed, MP and SP produce different cytokine profiles suggestive of cross-talk between the populations to maintain tumour cell growth. Furthermore, the PI 3-Kinase (PI3K) pathway was shown to be an important regulator of SP survival by regulating ATP-binding cassette (ABC) transporters hinting towards future therapeutic applications.

Due to the rarity of biopsy samples taken at relapse in paediatric B-NHL, resistance to chemotherapy was induced both *in vivo* and *in vitro* mimicking relapse. *In vivo*, this was achieved by administering multi-agent chemotherapy over an extended period to mice bearing established PDX tumours until apparent cure. The subsequent relapsed tumours contain a significantly increased proportion of SP cells and RNA sequencing identified PI3K and B cell receptor pathways as potential therapeutic targets in a subset of chemo-resistant tumours.

In order to identify further protein targets for the treatment of paediatric B-NHL that is resistant to standard chemotherapeutic agents including Rituximab, a CRISPR/Cas9-based genome-wide overexpression screen (SAM) was conducted with the human Burkitt lymphoma cell line Raji. The screen and subsequent validation highlighted several potential therapeutic targets including proteins in the PI3K, MAP kinase and Toll-Like receptor (TLR) signalling pathways, with different mechanisms of resistance identified between cells treated with standard chemotherapy drugs or Rituximab.

Overall, this thesis describes data showing that paediatric B-NHL consists of heterogeneous populations of tumour cells, including tumour initiating SP cells and MP progeny. The SP has significantly increased resistance to chemotherapy and supports the growth and survival of the MP through production of soluble factors with the PI3K pathway important in driving the SP phenotype. Furthermore, using both *in vivo* and *in vitro* models of resistance we have identified pathways that may be therapeutic targets in relapse/refractory disease.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the relevant Degree Committee.

Sorcha Forde

Declaration of received assistance

The following assistance was received from professional service providers

1. Bauer Core Facility (Harvard University, USA) - HiSeq-compatible libraries were prepared by adding Illumina adapters and barcodes via PCR amplification with custom oligonucleotides. The purified libraries were then submitted to Bauer Core for HiSeq high Output v4 reactions.
2. NGS library facility, Wellcome Trust, Medical Research Council Stem Cell Institute, University of Cambridge – Total RNA isolated from tumour samples was submitted for RNA-seq library preparation.
3. CRUK Cambridge Institute Genomics Core – prepared RNA-seq libraries were submitted for sequencing on one lane of an Illumina HiSeq 4000.
4. NIHR Cambridge BRC Cell Phenotyping Hub, Addenbrooke's Hospital, University of Cambridge – stained and prepared samples were submitted to the sorting facility for population sorting using a FACSAria™ Fusion.
5. Histology laboratory, Department of Pathology, University of Cambridge - Formalin fixed tumours were submitted to Helen Skelton for paraffin embedding and haematoxylin and eosin staining. Analysis of slides was performed by clinical histopathologist Dr. Olivier Giger at the Department of Pathology, Addenbrooke's hospital, University of Cambridge.
6. Cambridge haematopathology & oncology diagnostic service (HODS), Addenbrooke's Hospital - Tumour cell samples were prepared in Carbonyl's fixative and submitted to Bridget Manasse for FISH analysis.

All experiments in this thesis were completed by myself with the following assistance provided during the CRISPR screen -

1. 30 guide sequences were previously introduced into the lenti sgRNA(MS2)_Zeo backbone vector by a former member of the lab, Dr. Liam Lee.
2. Following amplification of the sgRNAs' guide sequence region and appending of the Illumina compatible adapters and barcodes, amplicons were then quantified by qPCR reactions by Dr. Liam Lee before submission to Bauer Core (Harvard University).

Acknowledgements

Firstly, I wish to thank my supervisor Dr. Suzanne Turner for her continuous support and guidance over the course of my PhD. I am grateful for Suzanne's scientific advice throughout different elements of my project but I also appreciate the opportunity she gave me to develop as an independent scientist. I am also thankful for the words of encouragement Suzanne always offered whenever I most needed them.

I wish to thank Dr. Amos Burke for his valuable clinical insights over the years as well as his assistance in obtaining patient samples. Most importantly I am grateful to the brave patients and their families for participating in the clinical research and consenting for biopsy samples to be used.

In particular, I wish to thank the Alex Hulme foundation for funding my PhD studentship and providing invaluable equipment for the lab. The strength and determination shown by Nic, Dave, Brenda and Matt is inspirational and I am proud to have been a part of the charity's work.

Thank you to my lab mates, especially to Liam Lee and Nina Prokoph for sharing their scientific knowledge with me. I am grateful for their friendship and scientific advice. Thank you to Stephen for his friendship and entertainment, Leila for assisting with injections and Gavin, Ricky and Hugo for helping whenever required. Thank you also to the staff in CBS for their excellent support in caring for and monitoring the mice throughout the 4 years.

Thank you to my siblings; Tríona, Neasa and Dinny as well as Graeme and my wonderful nephew Fionn. I'm very lucky to have such a supportive and encouraging family and I look forward to having more time to spend with you all now!

Thank you to my boyfriend PJ for your never-ending patience, kindness, support and good-humour over the past 4 years. It is not possible to express my gratitude in a few short words but I could not ask for a better partner and I am extremely grateful for all your help.

Finally, thank you to my mom and dad, Síle and Denis for your never-ending love, encouragement and kindness. I am truly grateful for both your moral and financial support throughout my entire education. Most importantly, thank you for nurturing my inquisitive mind and enabling me to reach my potential.

Summary

Paediatric mature B-cell non-Hodgkin lymphoma (B-NHL) is a highly aggressive disease with one of the fastest growth rates of any human malignancy. The subtypes are Burkitt and Burkitt-like lymphoma/leukaemia, diffuse large B-cell lymphoma and primary mediastinal B-cell lymphoma; with Burkitt lymphoma/leukaemia alone accounting for 38% of NHL diagnoses in patients under 15 years. Outcomes have improved significantly in recent years with the introduction of intensive multi-agent chemotherapy, followed by the addition of Rituximab, resulting in a survival rate exceeding 90%. Despite this, those patients who acquire resistance to chemotherapy and/or Rituximab have a dismal prognosis with few salvage therapy options. The aim of this thesis was to identify cell populations that are driving therapy resistance in order to develop targeted treatments for relapsed and therapy-refractory patients.

Utilising samples obtained from patients enrolled in the Inter-B-NHL ritux 2010 clinical trial, we established 5 paediatric Patient Derived Xenografts (PDXs) – to our knowledge the largest paediatric B-NHL resource of this type. These PDXs faithfully recapitulate tumour characteristics through passage and enable the identification of heterogeneous populations within these malignancies. One such population, known as the side population (SP), exhibits tumour initiating potential, expresses stemness-related genes and shows increased resistance to chemotherapy. *In vitro*, the SP produces soluble factors necessary for growth of the main population (MP) of cells (depleted of SP); indeed, MP and SP produce different cytokine profiles suggestive of cross-talk between the populations to maintain tumour cell growth. Furthermore, the PI 3-Kinase (PI3K) pathway was shown to be an important regulator of SP survival by regulating ATP-binding cassette (ABC) transporters hinting towards future therapeutic applications.

Due to the rarity of biopsy samples taken at relapse in paediatric B-NHL, resistance to chemotherapy was induced both *in vivo* and *in vitro* mimicking relapse. *In vivo*, this was achieved by administering multi-agent chemotherapy over an extended period to mice bearing established PDX tumours until apparent cure. The subsequent relapsed tumours contain a significantly increased proportion of SP cells and RNA sequencing

identified PI3K and B cell receptor pathways as potential therapeutic targets in a subset of chemo-resistant tumours.

In order to identify further protein targets for the treatment of paediatric B-NHL that is resistant to standard chemotherapeutic agents including Rituximab, a CRISPR/Cas9-based genome-wide overexpression screen (SAM) was conducted with the human Burkitt lymphoma cell line Raji. The screen and subsequent validation highlighted several potential therapeutic targets including proteins in the PI3K, MAP kinase and Toll-Like receptor (TLR) signalling pathways, with different mechanisms of resistance identified between cells treated with standard chemotherapy drugs or Rituximab.

Overall, this thesis describes data showing that paediatric B-NHL consists of heterogeneous populations of tumour cells, including tumour initiating SP cells and MP progeny. The SP has significantly increased resistance to chemotherapy and supports the growth and survival of the MP through production of soluble factors with the PI3K pathway important in driving the SP phenotype. Furthermore, using both *in vivo* and *in vitro* models of resistance we have identified pathways that may be therapeutic targets in relapse/refractory disease.

Abbreviations

ABC	ATP-binding cassette
ADCC	Antibody-dependent cellular cytotoxicity
AID	Activation induced cytidine deaminase
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
ANOVA	Analysis of variance
BCI	Bliss Combination Index
BCR	B cell receptor
BL	Burkitt lymphoma
B-NHL	B-cell non-Hodgkin lymphoma
Cas9	CRISPR-associated protein 9
CBS	Central biomedical services
CD	Cluster of differentiation
CLL	Chronic lymphocytic leukaemia
CLS	Categorical class
CMC	Cytokine mediated cytotoxicity
COG	Children's oncology group
COP	Cyclophosphamide, vincristine and prednisone
COPAD	Cyclophosphamide, vincristine, prednisone and doxorubicin
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR-based activation
CRISPRi	CRISPR-based inhibition
CSR	Class switch recombination
CSC	Cancer stem cell
DNA	Deoxyribonucleic acid
DGE	Differential gene expression
eBL	Endemic Burkitt lymphoma
EFS	Event free survival
EBV	Epstein Barr Virus
ED	Effective dose
EFS	Event free survival
ELDA	Extreme limiting dilution analysis
FACS	Florescent activated cell sorting
FBS	Fetal Bovine Serum
FD	Fold difference
FDR	False discovery rate
FFPE	Formalin fixed, paraffin embedded
FISH	Fluorescent in situ hybridisation
FL	Follicular lymphoma
FSC	Forward scatter
GC	Germinal centre
GCT	Gene cluster text
gDNA	Genomic DNA
GEP	Gene expression profiling
GO	Gene ontology

GOF	Gain of function
GPI	Glycosylphosphatidylinositol
GSEA	Gene set enrichment analysis
HIV	Human Immunodeficiency Virus
HODS	Hematopathology & oncology diagnostic service
HSC	Haematopoietic stem cells
HSF1	Heat shock transcription factor 1
Ig	Immunoglobulin
IP	Intraperitoneal
iPS	Induced pluripotent stem cells
ITH	Intra-tumour heterogeneity
IV	Intravenous
JAK- STAT	Janus kinase-signal transducer and activator of transcription
LIC	Leukaemia Initiation cells
LMB	Lymphomas Malins de Burkitt
LnRNA	Long non-coding RNA
MAGeCK	Model-based analysis of genome-wide CRISPR/Cas9 knockout
MAPK	Mitogen-activated protein kinase
MCL	Mantle cell lymphoma
MEK	MAPK/ERK kinase
miRNA	MicroRNA
MLE	Maximum likelihood estimation
MNC	Mononuclear cell
MOI	Multiplicity of infection
MP	Main population
MRP	Multidrug resistance-associated protein
MSigDB	Molecular signatures database
mTOR	Molecular target of rapamycin
MZL	Marginal zone lymphoma
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NHL	Non-Hodgkin lymphoma
NK	Natural killer
NOD	Non-obese diabetic
NSG	NOD scid gamma
NT	Non-targeting
OS	Overall survival
OV	Overexpression
PAM	Protospacer adjacent motif
pBL	Paediatric Burkitt lymphoma
pB-NHL	Paediatric B-cell non-Hodgkin lymphoma
PBS	Phosphate-buffered saline
PCA	Principal component analysis
PDX	Patient-derived xenograft
PFS	Progression free survival
PI	Propidium iodide
PI3K	Phosphoinositide 3-kinase
PBMC	Peripheral blood mononuclear cell

PMLBL	Primary mediastinal large B-cell lymphoma
PRECOG	Prediction of clinical outcomes for genomic profiles
QC	Quality control
RIGER	RNAi gene enrichment ranking
RIN	RNA integrity number
RNA	Ribonucleic acid
RNAi	RNA interference
RPMI	Roswell Park Memorial Institute
RRA	Robust ranking aggregation
SAM	Synergistic activation mediator
sBL	Sporadic Burkitt lymphoma
SC	Subcutaneous
SD	Standard deviation
sgRNA	Single guide RNA
SHH	Sonic hedgehog
SHM	Somatic hypermutation
SP	Side population
SSC	Side scatter
TA	Transient-amplifying
TCL	T-cell lymphoma
TFG-β	Transforming growth factor-β
TIC	Tumour initiating cell
TLR	Toll-like receptor
TME	Tumour microenvironment
TPC	Tumour propagating cells
TSS	Transcription start site
UV	Ultra-violent
WHO	World Health Organisation

Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Lymphoma

Lymphomas are a highly heterogeneous group of cancers that are characterised by neoplastic transformation of lymphocytes at specific stages in lymphocyte ontogeny. Due to the diversity of this group of malignancies, the field is ever-evolving and in the 4th version of the World Health Organisation (WHO) classifications, which was updated in 2016 there are now greater than 70 distinct forms of this disease. These are classified broadly based on cell of origin; initially distinguishing B cell, T cell and natural killer (NK) cell disorders, with subsequent characterization based on the stage of lymphocyte ontogeny as well as morphology, molecular profile and clinical characteristics[1]. Lymphoma is a relatively rare cancer with just under 16,000 cases diagnosed in the UK in 2015[2]. These cases are classified clearly as either Hodgkin or non-Hodgkin lymphoma (NHL) with the former, characterised by the presence of Reed–Sternberg cells, accounting for only 10-14% of all diagnoses[2]. B-cell non-Hodgkin lymphoma (B-NHL) represents the most common subtype of NHL in childhood and adolescence, accounting for up to 60% of all newly diagnosed lymphomas[3,4]. In contrast to adults, NHL in children and adolescents are almost exclusively highly aggressive. Most paediatric B-NHLs (pB-NHL) are Burkitt lymphoma/leukaemia (BL/B-AL) and diffuse large B-cell lymphoma (DLBCL), whereas only a small percentage are represented by primary mediastinal large B-cell lymphoma (PMLBL), follicular lymphoma (FL), nodal marginal zone lymphoma and intermediate lymphoma[5]. BL alone accounts for almost two-thirds of B-NHL in children and adolescents but the incidence varies with age as outlined in Figure 1.1[6]. BL is the predominant form of NHL in 0-14 year old children but in the 15-19 adolescent age category DLBCL becomes more prominent. This age-related reduction in BL incidence continues in adulthood with only approximately 5% of NHL diagnosed as BL above the age of 20. DLBCL is one of the most common forms of NHL in adults accounting for 30% of cases in the greater than 20 years age category.

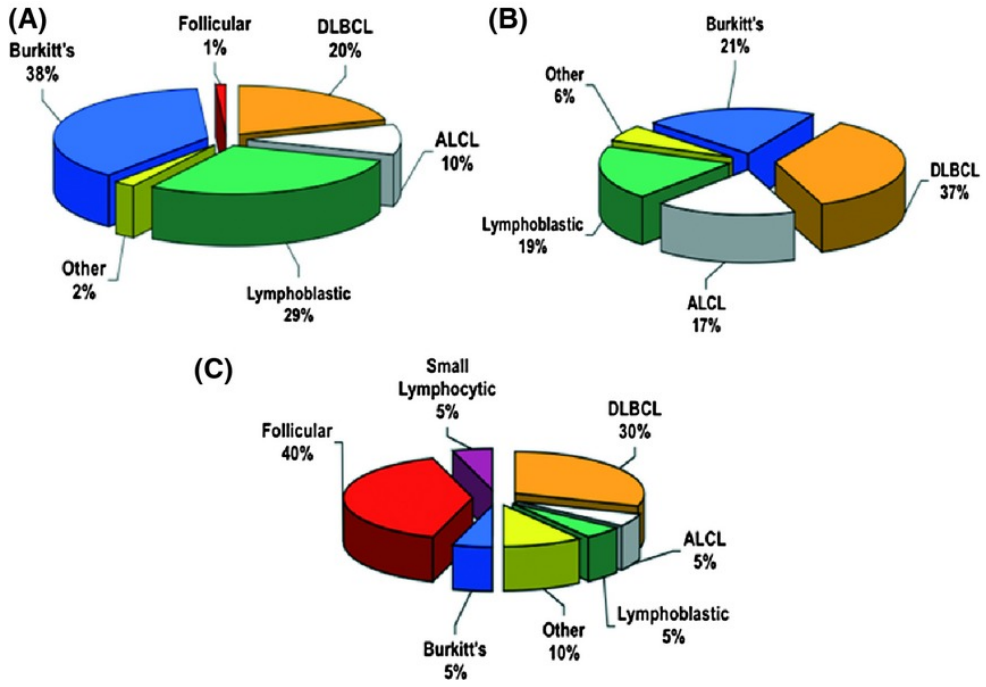


Figure 1.1. Incidence of non-Hodgkin lymphoma subtypes in (A) 0-14 years age group, (B) 15-19 years age group, and (C) 20 years and older age group.[6]

1.2 B cell lymphomagenesis

A detailed understanding of the development processes involved in healthy B cells has been paramount in elucidating potential stages when oncogenic perturbations may occur. The end result of B cell development is the generation of a large pool of cells with a broad range of antigen recognition, capable of producing antigen-specific immunoglobulins (Ig)[7]. The processes involved in producing high affinity B cells include V(D)J recombination, somatic hypermutation (SHM) and class switch recombination (CSR). The involvement of these mechanisms in B cell lymphomagenesis is evidenced by the fact that approximately 95% of lymphomas are derived from B cells despite there being a similar proportion of T and B cells in the human body[8,9]. B cell development begins in the bone marrow with the rearrangement of the heavy and light chain genes in the process of V(D)J recombination to produce a B cell receptor (BCR) on a committed

B cell. The cells pass several selection processes and checkpoints before entering the blood as mature, naïve B cells[7,10,11]. Naïve B cells then accumulate in secondary lymphoid organs where antigen-dependent development initiates and further risk of DNA damage and translocation emerge. Following antigen engagement, mature B cells become activated and enter T cell rich areas of the lymphoid organs to generate germinal centres (GCs). Here, the cells undergo clonal expansion as well as the processes of SHM and CSR which take place in the dark and light zones, respectively, and result in further remodelling of the Ig loci[12]. SHM increases antigen affinity while B cells undergo CSR to produce alternative isotypes of antibody by rearrangement of the constant region of the Ig heavy chain. Cells in the light zone are also subject to selection based on affinity of their BCR and will either progress into plasma or memory cells, or alternatively be returned to the dark zone for further specification of Ig genes[13]. A key regulator of the CSR and SHM processes is the enzyme activation induced cytidine deaminase (AID) which has been implicated in lymphomagenesis. Its activity is not specific to the Ig locus and hence there is a risk of mutations occurring in oncogenes as well as DNA strand breaks resulting in instability and translocations[14].

Studies have highlighted that the status of the Ig genetic regions within lymphoma cells can pinpoint their origin to a specific stage of B cell development. A large number of B cell lymphomas appear to be GC educated, evidenced by the presence of hypermutated V-regions in their Ig loci[15]. Subtypes can be further assigned to specific compartments within the GC owing to their tendency to adopt the gene expression program of their normal GC B cell counterparts[16,17,18,19]. Indeed, the maintenance of some components of normal B cell physiology is likely to be important for survival of B-cell lymphomas and translocations predominantly occur in the non-productively rearranged Ig locus, rather than the allele that produces the productive Ig heavy chain that is required to construct the BCR[20].

1.3 Burkitt Lymphoma

BL is composed of monomorphic, intermediate-sized lymphocytes with a GC B cell immunophenotype. Cells are positive for the pan-B-cell antigens CD20, CD79a and PAX5 as well as GC antigens CD10 and BCL6, but are negative for the anti-apoptotic protein BCL2. BL is the most aggressive tumour known with a turn-over rate of approximately 24 hours[21]. It was first discovered by Irish surgeon Denis Burkitt when he described 38 cases of "sarcomas" involving the jaws of African children at Mulago Hospital in Kampala Uganda, providing the first clinical description of BL[22]. BL has played a fundamental role in our understanding of tumourigenesis over the years. In collaboration with Burkitt, virologists Epstein and Barr documented the first case of a tumour-associated virus when they extracted viral particles of the Epstein Barr virus (EBV) from a sample of endemic BL in 1964[23]. BL was also pivotal in elucidating the role of MYC in cancer and led to the discovery of the first oncogenic chromosomal translocation[24–26].

1.3.1 Epidemiology and aetiology of BL

There are three distinct subtypes of BL – sporadic BL (sBL), endemic BL (eBL) and immunodeficiency associated BL. Although they display unique geographical patterns and have slightly different clinical features, they have common morphology and immunophenotype, and genetic profiling highlights few differences between the entities[27–30].

- Endemic BL

After discovering BL, Burkitt mapped the incidence across sub-Saharan Africa and noticed a trend whereby prevalence of BL and endemic malarial infection overlapped. Further studies also elucidated that the incidence of BL correlated with the intensity of malarial infection and not just geographical distribution[31,32]. Despite strong evidence for its potential causative role in BL, the exact role of malaria in tumourigenesis has not been established. The ability of malaria to drive chronic B cell activation has been proposed as one possible mechanism as this may predispose B cells to the acquisition of IG/MYC translocations[33]. A second potential mechanism involves EBV and proposes that exposure to malarial antigens during multiple episodes of infection could reactivate EBV

in memory B cells where it had been latent, thereby increasing viral load and the number of EBV-infected cells[34,35]. Like malaria, the role of EBV in eBL is also elusive, despite being present in over 95% of cases[36,37]. Indeed because it is present in only a minority of sBL cases it cannot be deemed necessary for pathogenesis. EBV induces the proliferation of resting human B cells, leading to the expression of a variety of EBV-encoded latency proteins, many of which modulate key regulatory pathways such as Phosphoinositide 3-kinase (PI3K) and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB)[38]. The current understanding therefore, is a 'hit-and-run' role for EBV in carcinogenesis. By inhibiting the apoptosis of premalignant tumour cells, EBV facilitates the occurrence of transforming events following which the majority of its latency proteins are no longer required[39].

- Sporadic BL

Soon after the clinical description of BL was established in Africa, pathologists in Europe and the USA described lymphomas that were histologically identical to the African entity. It was therefore realised that this disease was not exclusive to Africa, although the incidence there is markedly higher. The most significant difference between the two forms is that EBV is only associated with around 20% of non-endemic forms[39]. The unknown aetiology of the malignancy led to the name 'sporadic' BL. Other interesting clinical characteristics of sBL are that it appears to have a bimodal age-related incidence in males with a peak in childhood and again at above 60 years of age. sBL is also predominantly a male disease with one study of 80 paediatric patients having 89% males and a second study of 67 patients giving a slightly lower male dominance at 67%[40,41]. BL is considered to be GC-derived, although the origin of the c-MYC translocation has not been fully elucidated in BL and is suggested to be different in sBL and eBL. The translocation present in eBL seems to be acquired due to aberrant SHM in early GC-phases whereas the translocation mechanism in sBL is thought to be faulty CSR in the GC light zone, prior to re-entry into the GC dark zone[13,27,39].

- Immunodeficiency-related BL

Human immunodeficiency virus (HIV) patients were observed to be predisposed to NHL, with BL accounting for 10-40% of cases[42]. This led to the inclusion of a third BL variant in the WHO classification of haematological malignancies known as immunodeficiency-related BL. Although an increased risk of BL exists in other immunocompromised patients, it is not as great as in HIV-infected individuals[43]. Approximately 30-40% of immunodeficiency-related BL cases are associated with EBV, and immunosuppression is believed to contribute to its genesis by conferring an increased EBV load and deregulating immune mechanisms[44–46].

1.3.2 Molecular background to BL

1.3.2.1 C-MYC translocation

All forms of BL are invariably associated with chromosomal translocations that dysregulate the expression of c-*MYC*, a gene encoding a basic helix-loop-helix transcription factor that binds to DNA in a sequence-specific fashion. MYC is a global transcription factor that normally plays a central role in the transcriptional regulation of an emerging set of downstream genes that control diverse cellular processes. It is estimated to govern approximately 10-15 % of genes in the genome, influencing several aspects of survival and proliferation of cells including cell cycle progression, differentiation, metabolism and apoptosis[26,47]. The chromosomal translocations characteristic of BL always result in the juxtaposition of the DNA coding sequences for c-*MYC* with sequences from Ig gene enhancers. Ig enhancer elements bind to B cell–specific factors capable of activating transcription from genes located up to 500,000 base pairs away. Due to the fact that Ig enhancer elements are specifically active in mature B cells, their juxtaposition to c-*MYC* in BL cells drives inappropriately high levels of c-*MYC* mRNA and protein expression[48,49]. In 80% of cases of BL, the translocation partner for c-*MYC* is the IgH locus, resulting in the formation of a t(8;14)(q24;q32) aberrant chromosome. These translocations typically occur on the background of a simple karyotype lacking additional rearrangements[39]. The remaining 20% of BL have alternative variants of MYC translocations, with the MYC gene partnered with the κ locus

on chromosome 2 giving the t(2;8) or the light chain locus on chromosome 22 leading to the t(8;22) translocation[47].

1.3.2.2 Other genetic features of BL

The detection of the t(8;14) translocation in blood cells of healthy individuals highlights that this translocation alone is not sufficient for lymphomagensis[50]. This was further supported by the first mouse model of BL. This model of deregulated MYC mimicked the t(8;14) chromosomal translocation by bringing the MYC gene under the control of the Ig heavy-chain intronic enhancer (Eμ)[51]. However, these Eμ/MYC transgenic mice are not faithful models for human BL as the tumours that develop more closely resemble pre-B or naïve B cells, not GC B cells.

Independent studies by Love *et al.,* Richter *et al*., and Schmitz *et al.* using high throughput sequencing approaches provided an overview of the genomic landscape of BL and gave greater insights to the molecular basis of BL[16,17,52]. The most common mutation in BL is in *MYC* itself, present in 70% of cases and often affecting its transactivating domain. As deregulated MYC is already present due to the translocation these mutations often involve the uncoupling of the pro-proliferative mechanisms of MYC from its ability to induce apoptosis by impairing transactivation of the pro-apoptotic protein BIM[53]. *TP53* mutations are reported in up to 35% of BL cases and also likely function to counterbalance the pro-apoptotic effects of MYC while identified *GNA13* mutations may be involved in GC B cell migration[54,55]. Novel mutations in the Inhibitor of DNA binding 3 (ID3) gene were reported in all three studies but shown to vary in incidence between 34% (Love *et al.*), 58% (Schmitz *et al*.) and 68% (Richter *et al*.) ID3 inhibits the transcription factor TCF3 and Schmitz *et al.* showed that TCF3 itself is also frequently mutated in BL. Either inactivating mutations in *ID3* or activating mutations in *TCF3* result in increased expression of downstream targets of TCF3 leading to increased cell proliferation and survival via PI3K[16,17,52]. Interestingly, *Cyclin D3* (*CCND3*), a direct target of TCF3, was also found to have mutations in 38% of cases reported by Richter *et al*[16]. Due to the discrepancy in mutation rates in these studies Rohde *et al.* undertook a study of 84 uniformly diagnosed paediatric patients and

determined a mutation frequency of 78% (*ID3*), 13% (*TCF3*) and 36% (*CCND3*) with a total of 88% of patients having a mutation in at least one of these genes[56]. Taken together mutations in this important axis may provide the vital cooperating mutations required for the pathogenesis of BL.

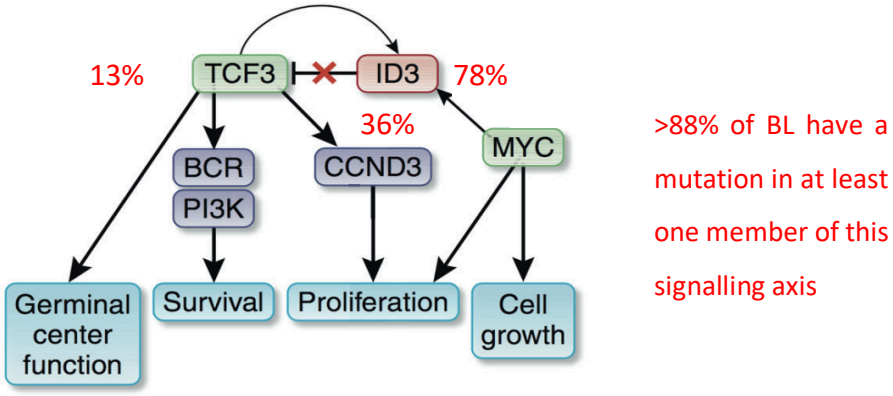


Figure 1.2. The ID3-TCF3-CCND3 oncogenic pathway in BL.[56]

1.3.3 Signalling Pathways in BL

The urgent need to develop less toxic therapies as well as more targeted strategies to overcome chemo-resistant disease requires a detailed understanding of oncogenic signalling pathways in pBL. A study by Kuppers *et al.* in 2005 provided the first insights into the potentially important role of BCR signalling in BL[15]. They showed the MYC translocation in BL occurs only to the non-productively rearranged Ig heavy-chain locus, and not the allele that produces the productive Ig heavy chain that is required to construct the BCR. This suggests that BL may rely on BCR signalling for survival. BCR signalling takes two forms – tonic BCR signalling which activates PI3K signalling, and active BCR signalling which engages multiple pathways including NF-κB in an antigen-dependent manner[57]. Unlike other aggressive B cell lymphomas, BL does not show strong expression of genes downstream of NF-κB suggesting that this latter method of BCR signalling is not as important for BL pathogenesis[58]. The fact that mice with

constitutively active PI3K in conjunction with MYC expression produce tumours resembling BL at the histological and gene expression level further implicates this signalling pathway in tumourigenesis[59]. Interestingly, tertiary mutations acquired in these murine tumours include those in *CCND3* which also occurs in human BL, highlighting the validity of this model.

Evidence of the preferential role of tonic BCR signalling in BL is provided by studies conducted with cell lines. In 67% of cases, knockdown of BCR subunits or the BCR-associated tyrosine kinase SYK, resulted in decreased PI3K activity inducing apoptosis, whereas inhibition of NF-κB had no impact on cell viability[17]. The importance of the TCF3-ID3 axis in PI3K signalling is also explained by the presence of TCF3 at the Ig heavy-chain 3' enhancer, up-regulating expression of the Ig heavy and light chains and increasing expression of the BCR on the cell surface[17]. Further mechanisms involved in augmenting PI3K signalling in BL include the presence of *PTEN* inactivating mutations in approximately 7% of cases[39]. The induction of expression of the *mir*-17-92 cluster by MYC also reduces expression of PTEN via Mir-19a and mir-19b[60,61]. This evidence, together with the knowledge that gene expression signatures of PI3K activity are more highly expressed in human BL biopsies than in other lymphoma subtypes, strongly supports the important role of PI3K signalling in BL[17,59].

A recent study in pBL identified potential new pathways involved in lymphomagenesis; Lee *et al.* analysed the comparative genomic signature and targetable signalling pathways in pBL samples from the Children's Oncology Group (COG) ANHL01P1 study by genomic profiling and compared the results to pBL samples from public databases[62]. This study led to some interesting insights – data from three completely independent databases identified significant overexpression of genes involved in Toll-like receptor (TLR), Janus kinase/signal transducers and activators of transcription (JAK-STAT) and mitogen-activated protein kinase (MAPK) signalling. Several of the genes identified in these pathways offer attractive therapeutic targets and warrant further investigation.

1.3.4 Prognostic indicators in BL

Currently, very few genetic prognostic indicators have been determined in pBL and staging is based on factors such as bone marrow and CNS involvement as well as lactate dehydrogenase levels. Paediatric B-NHL is curable in up to 90% of cases, yet those patients who relapse have only a 20-30% chance of survival[63]. For this reason, it is imperative that prognostic indicators are identified to better stratify patients and direct the course of therapy. Early studies using conventional cytogenetics and fluorescent in situ hybridisation (FISH) identified recurrent genetic abnormalities in BL, including gain of 1q and 13q and loss of 17p, and found that 13q abnormalities are associated with a worse prognosis[64,65]. Using more advanced techniques, studies validated the significance of the 1q and 17p abnormalities although these studies were not specific to pBL[66–68]. A more recent study utilised a genome-wide, high-resolution molecular inversion probe SNP assay to identify copy number gain alterations in DNA isolated from FFPE pBL diagnostic biopsy tissues[69]. Amplification of 13q, 7q gains and copy-neutral loss of heterozygosity of 5q were identified as being more common in pBL and although the sample size limited statistical power, patients with a 13q31 gain were more likely to relapse over a minimum follow-up period of 6.5 months. This region of chromosome 13 includes 3 genes and 8 miRNAs: *GPC5, GPC6, LINC00410, MIR622*, and the *MIR17HG* gene encoding the MIR17-92 cassette of 7 miRNAs. The *MIR17HG* cluster on chromosome 13 is oncogenic and has been shown to accelerate lymphoma development in a mouse model and increase tumourigenicity in cooperation with MYC[70,71]. MIR17HG has been established as a MYC transactivation target in a B cell line model and its expression correlates positively with MYC expression in colorectal tumours[72,73]. Recently it was shown to regulate expression of genes involved in breast cancer development and progression[74]. Interestingly, in this study Schiffman *et al.* also identified a strong correlation between 13q31 gain and loss of 17p13 which includes the *TP53* locus[69]. MIR17HG induces anti-apoptotic effects via PTEN and E2F1 and it is possible that this could synergise with loss of TP53 to confer a strong proliferative program and result in a subset with an increased aggressive phenotype[75].

1.4 DLBCL

The DLBCLs are mature B-cell neoplasms characterized by expression of surface Ig and B cell–associated surface markers (CD19, CD20, CD22, and CD79a)[76]. According to gene expression profiling (GEP) classification, 2 main molecular subgroups exist with differing cells of origin and clinical outcome - Germinal centre B-cell (GCB) and activated B-cell (ABC) DLBCL[77]. It is important to note that the GCB subtype is molecularly more similar to BL and accounts for the vast majority of paediatric cases (age 0-9 years 95% GCB, 10-19 years 80% GCB, adults 42% GCB)[78,79]. GCB DLBCL has also been consistently associated with improved outcomes over non-GCB disease[78,80]. Common oncogenic signatures in both subtypes include genetic aberrations in genes that overcome BCL6 regulation such as *B2M*, as well as lesions in chromatin modifiers affecting epigenetic regulation (*CREBBP, EP300, MLL2*)[13,81]. Frequent mutations in the ABC subtype include those in *MYD88, KLHL14, CD79B* and *SIGLEC10* which result in cell survival via chronic BCR signalling and activates NF-κB while GC-DLBCL is associated with mutations in *GNA13*, *BCL2* and *EZH2*[82,83]. A t(14;18)(q32;q21) translocation of BCL2 is present in 35% of adult but is uncommon in paediatric cases[84]. Similar to BL, PI3K has been implicated in GC-DLBCL with deletion of *PTEN* found in 11% of cases while amplification of *miR-17-92* is seen in approximately 13% of cases[81,85]. *IRF4* oncogene rearrangements, which have been identified in a small subset of B-NHLs with a distinct immunophenotype from the GCB and ABC subtypes of DLBCL, occur more commonly in those younger than 18 years of age[86]. GEP analysis by Dave *et al*. compared BL and DLBCL and established a reliable method to distinguish the two entities at the molecular level[58]; MYC and its target genes as well as a subgroup of GC B cell genes were more highly expressed in BL. DLBCL on the other hand, display higher expression of (major histocompatibility complex) MHC class 1 genes and NF-κB target genes. DLBCL cases were also effectively grouped into the ABC and GCB subtypes[58]. This research was carried out in adult case studies although paediatric cases produced similar findings. Klapper *et al.* reassigned diagnoses according to molecular signatures and compared results to the original morphological diagnosis in paediatric cases[87]. Morphological diagnoses of BL were confirmed by molecular methods but 31% of DLBCL cases were re-classified as BL following molecular analysis. Many of the re-classified cases had interesting features

which may have initially excluded them from a BL diagnosis, such as BCL2 expression or a below average ki67 index. Between 5 and 10% of paediatric DLBCL cases can carry a MYC translocation which can further complicate diagnosis[88].

1.5 Current treatment regimens for paediatric B-NHL

1.5.1 Chemotherapeutic backbone

The past three decades have seen a dramatic increase in survival for pB-NHL and event free survival (EFS) rates have almost doubled from the late 1970s to the present era[21]. Consecutive prospective trials, such as the French Society of Pediatric Oncology and French-American-British/Lymphomes Malins de Burkitt (FAB/LMB), and the oligo-national Berlin-Frankfurt-Munster (BFM) studies have resulted in significant treatment improvements. These focused on risk-stratification to minimise acute and long-term toxicity for patients with a favourable prognosis and to intensify treatment for high risk patients[89,90].

The FAB/LMB backbone chemotherapy protocols currently used in the UK are based upon three risk stratification groups (Figure 1.3). Group A, low risk patients, are treated with two cycles of cyclophosphamide, vincristine, prednisone and doxorubicin (COPAD) following resection. This yields excellent outcomes, with a 98.3% 4-year EFS[91]. Intermediate-risk patients in group B receive backbone chemotherapy consisting of low-dose cyclophosphamide, vincristine, and prednisone (COP) followed by two induction cycles (COPAD plus methotrexate) and two consolidation cycles (cytarabine plus methotrexate), each with intrathecal chemotherapy given throughout, resulting in an EFS rate of 91%[90]. The highest risk patients in group C, achieve 79% EFS with the FAB/LMB backbone, strengthened by the addition of high-dose methotrexate during induction and cytarabine plus etoposide for two cycles of intensification followed by four cycles of maintenance chemotherapy[89].

The BFM chemotherapy backbone is similar to the FAB/LMB regimen, with the main difference being the stratification of patients into four groups and the incorporation of

ifosfamide into the regimens, compensating for lower doses of doxorubicin and cyclophosphamide.

	Berlin-Frankfurt-Munster	French-American-British
Low risk ↓ High risk	**R1** Stage I or II, completely resected	**Group A** Resected stage I and abdominal completely resected stage II
	R2 Stage I or II, not resected Stage III with LDH <500 U/L	**Group B** All patients not in Group A or C
	R3 Stage III with LDH ≥500 to <1000 U/L Stage IV with LDH <1000 U/L and CNS-negative	
	R4 Stage III or IV with LDH ≥1000 U/l and/or CNS-positive	**Group C** Bone marrow disease (≥25% L3 blasts) and/or CNS-positive

Figure 1.3. Risk stratification groups for paediatric B-NHL[92]

	BFM	FAB	BFM	FAB	BFM	BFM	FAB
Group	**R1**	**Group A**	**R2**	**Group B**	**R3**	**R4**	**Group C**
Definition	Resected	Resected	Not resected, I, II, III, LDH <500	Not resected, I, II, III, IV, CNS-neg	III, LDH 500-999, IV plus B-ALL, LDH <1000, CNS-neg	LDH >1000 and/or CNS-pos	B-ALL IV, CNS-pos
No. of courses	2	2	4	4	5	6	8
MTX g/m², infusion	1, 4 h, × 2	—	1, 4 h, × 4	3, 3 h, × 4	5, 24 h, × 4	5, 24 h, × 4	8, 4 h (CNS-pos; 24 h), × 3 (4)
Dox mg/m²	50	120	100	120	100	100	240
CP g/m²	1.4	3	2.4	3.3	2.4	2.4	6.8
Ifo g/m²	4	—	8	—	8	8	8
Eto mg/m²	200	—	400	—	900	1400	2500

Table 1.1. Comparison of chemotherapeutic regimens and doses between FAB and BFM treatment protocols for pB-NHL.[93]

1.5.2 Introduction of Rituximab to paediatric B-NHL treatment

In order to try to improve upon the survival rates in pB-NHL and to reduce relapse, the anti-CD20 antibody Rituximab was evaluated in combination with chemotherapy in the Inter-B-NHL Ritux clinical trial. The rationale for this study was the success of combining

Rituximab with multi-agent chemotherapy in adult DLBCL and BL, as well as some anecdotal evidence of efficacy in paediatric cases[94–101].

The FAB/LMB chemotherapy back bone strategy was chosen for this study with some minor modifications. Rituximab was included at the start of treatment with two administrations at 48 hour intervals at the beginning of the two first courses of chemotherapy, and one administration at the beginning of the two following courses of chemotherapy. An interim analysis highlighted exceptional response in the Rituximab arm with a 1-year EFS rate of 94.2% in the Rituximab arm and 81.5% in the control arm[102]. Randomisation was halted and the trial completed with all enrolled patients receiving Rituximab.

1.6 Relapsed and refractory paediatric B-NHL

Despite significant improvements in overall curative rate brought about by the BFM and FAB/LMB trials as well as the success of Rituximab, relapse and refractory disease still occur. This is the greatest challenge in the current era of pB-NHL and salvage of these patients has remained extremely difficult. Several studies have put the overall survival (OS) probability in relapsed and refractory disease at less than 30%[21,89]. Rituximab has proven effective at reducing the probability of relapse, although, it is not yet known whether it will be effective at salvaging those cases that do relapse. One Japanese study showed improvement in prognosis of relapse or refractory pB-NHL in the Rituximab era but this has not yet been reported in other studies[103].

Based on the dismal salvage rates for patients who experience relapse or refractory disease, the urgent requirement for improved targeted therapies is obvious. In addition, the significant acute and long-term toxicities associated particularly with the intensity of group C FAB/LMB chemotherapy, must be considered. Ideally, targeted therapies would be effective in the treatment of relapsed and refractory cases, but might also be combined with reduced chemotherapy enabling dose de-escalation in upfront treatment of low risk patients. Due to both the rarity of occurrence and often the lack of biopsies taken of relapsed disease, few studies have been carried out on relapsed and refractory pB-NHL. For this reason, beyond the potential prognostic indictors mentioned

earlier, little is known about the cause of resistance and it is imperative that this becomes the focus of studies if we are to further improve the treatment of this disease. Studies in adult B-NHL are paving the way for potential new treatments including humoral and cellular immunotherapies targeting CD20 and CD19[104–107]. Targeted therapies that interfere with the signalling pathways driving lymphomagenesis are also likely an important area of novel therapeutics. The BTK inhibitor Ibrutinib has shown extraordinary potential in adults with chronic lymphocytic leukaemia (CLL), mantle cell lymphoma (MCL), and DLBCL[108–111] and a trial comparing the efficacy of Ibrutinib to Rituximab in paediatric and young adult participants with relapsed or refractory mature B-NHL is underway. The PI3K inhibitor Idelalisib has shown an objective response in phase II data evaluating adults with relapsed indolent mature B-NHL, as well as significantly improved outcomes in combination with Rituximab in adults with relapsed CLL[112]. Interestingly a phase 1b trial has recently opened investigating Idelalisib in combination with Rituximab and chemotherapy in children and adolescents with relapsed or refractory DLBCL or mediastinal B-cell lymphoma.

Overall, potential for novel therapies exists, however we must endeavour to gain a better understanding of pB-NHL biology, particularly in relapse, in order to identify therapies that will improve outcomes for patients who fail to respond to conventional methods.

1.7 Mechanisms of action and resistance to current therapeutics in paediatric B-NHL

1.7.1 Chemotherapeutic agents

The main chemotherapeutics common to both the FAB/LMB and BFM regimens and utilised during the induction stages of treatment are cyclophosphamide, vincristine, prednisone, doxorubicin and methotrexate. Each of these chemotherapeutic agents have diverse mechanisms of action leading to cell death.

Doxorubicin is an anthracycline and two alternative mechanism of actions have been proposed[113]: The first involves intercalation of the agent into DNA and disruption of topoisomerase II-mediated DNA repair resulting in DNA damage and cell death[114]. The main methods of resistance to doxorubicin identified involve increased expression of TOP2A, often by amplification, as well as enhanced drug efflux mediated by a number of ATP-binding cassette (ABC) drug transporters including ABCB1 and ABCC1 as well as ABCC2, ABCC3, ABCG2 and RALBP1[115–118]. The second mechanism of action of doxorubicin involves the generation of free radicals and subsequent damage to cellular membranes, DNA and proteins[119]. Doxorubicin is oxidized to semiquinone, an unstable metabolite, which is converted back to doxorubicin in a process that releases reactive oxygen species and triggers apoptotic pathways of cell death[120]. A host of emerging resistance mechanisms to doxorubicin have been reported in the literature. In T-ALL, α2β1 integrin was shown to promote chemoresistance to doxorubicin in cancer cells through MAPK/ERK signalling while in BL cell line studies, increased activation of PI3K/AKT was shown to mediate resistance to this drug[121,122].

Methotrexate, an antifolate, is a potent inhibitor of dihydrofolate reductase, the rate-limiting enzyme in the production of tetrahydrofolate. By decreasing the production of purines and pyrimidines, which are critical to *de novo* synthesis of DNA in mammalian cells, methotrexate is highly effective at halting cell proliferation and turnover[123]. A number of resistance mechanisms have been elucidated for methotrexate involving impaired cellular influx or increased efflux, increased expression, mutation of cellular targets, or intra-cellular accumulation of tetrahydrofolate cofactors. In methotrexate-resistant cell line models, both mutations of the *Reduced Folate Carrier* gene and

overexpression of the protein have been reported, and a deactivating mutation in this gene is associated with a worse clinical outcome in children with acute lymphoblastic leukaemia (ALL) receiving methotrexate[124–126]. Efflux mechanisms have also been implicated in resistance to antifolates including ABCB1, ABCG2 and ABCC 1-4 while amplification of the DHFR gene itself was reported in ALL and ovarian cancer, but the clinical relevance was not fully elucidated[127–133].

Vincristine is a member of the vinca alkaloids which inhibit the assembly of microtubules by binding to tubulin. This results in dissolution of the mitotic spindle necessary for mitosis, inducing apoptosis via loss of BCL2 function together with elevation of p53 and cyclin-dependent kinase inhibitor 1A expression[134]. Like methotrexate and doxorubicin, vincristine is a substrate for ABC transporters and ABCB1, ABCC1-3, ABCC10 and RALBP1 have been reported in association with vincristine resistance[135–137]. More recent studies have also begun to highlight the potential role of targetable pathways in mediating vincristine resistance. In ALL, resistance to vincristine was shown to be mediated by the PI3K-Akt pathway while MAPK signalling increased resistance to vincristine in gastric cancer cells by regulating ABCG2[138,139].

Cyclophosphamide is the precursor of an alkylating nitrogen mustard antineoplastic used across a wide range of tumour types in childhood cancer. In order for cyclophosphamide to exert its antitumor activity, the prodrug requires metabolic activation by hepatic cytochrome P-450 enzymes to generate active alkylating species[140]. Alkylating agents have three main mechanisms of DNA damage; 1. Alkyl groups attach to DNA bases, resulting in DNA being fragmented by repair enzymes and preventing DNA synthesis and RNA transcription at the affected site. 2. They produce DNA damage by the formation of cross-links which prevents DNA from being separated for synthesis or transcription, and 3. They induce mispairing of the nucleotides leading to mutations[141]. Interpatient variability in the pharmacokinetics of this drug is thought to have a significant impact on efficacy of cyclophosphamide while several methods of resistance to alkylating agents have also been reported involving the DNA repair enzyme MGMT, the MMR pathway and de-regulation of apoptosis genes including *TP53* and

BCL2[142,143]. The role of ABC transporter efflux in response to cyclophosphamide is less clear but it has been reported to be a substrate of ABCB1[144].

1.7.1.1 Multi Drug Resistance

The evidence of ABC transporter involvement in mediating resistance to each of these chemotherapeutic agents suggests that they may be important in multi drug resistance in pB-NHL, although their clinical relevance remains ambiguous. P-glycoprotein, also called MDR1, is encoded by the *ABCB1* gene and was the first identified drug efflux transporter over 40 years ago[145]. Several years later, multidrug resistance-associated protein 1 (MRP1), encoded for by the *ABCC1* gene was identified and shown in cell line models to mediate resistance to doxorubicin, etoposide and vincristine among others[146]. Three independent groups then identified breast cancer resistance protein, BCRP1, encoded by *ABCG2*, and interest in this area increased[147–149]. First-generation inhibitors, focusing on MDR1, including verapamil, quinidine, amiodarone and cyclosporine A, were identified and added to chemotherapy regimens in clinical trials[150]. However, these agents lacked potency and had high toxicity with little evidence of inhibition of MDR1 in patients[151]. Second generation inhibitors also had poor success[152]. Third generation inhibitors were more potent with fewer pharmacokinetic interactions although toxicity linked to inhibition of MDR1 in normal tissues was reported[153]. Some third generation inhibitors showed cross reactivity with MRP1 and/or ABCG2 and in Acute Myeloid Leukaemia (AML) combination of cyclosporine A with chemotherapy led to increased response in one study[154]. However, despite a few successes the majority of clinical trials did not confirm clinical benefit and interest in the area waned[153]. Several decades after their initial discovery and with the emergence of new data and advanced techniques, it has been proposed that these transporters may be worth re-evaluating clinically. Recently, attention has been paid to the use of drug delivery systems, such as nanoparticles, primarily to restrict ABC transporter inhibition to tumour cells and reduce the non-selective inhibition in non-target organs[155]. Furthermore, identification of pathways and oncogenic kinases in regulating ABC transporter activity also offers promising potential for targeting these efflux pumps.

1.7.2 Rituximab

Unlike traditional chemotherapeutic agents which target all dividing cells, Rituximab is an anti-CD20 monoclonal antibody. The CD20 receptor appears on B cells during the pre-B-cell stage and is not expressed in other tissues, making Rituximab a targeted therapy and the first antibody-based treatment for cancer which received FDA approval[156]. A number of mechanisms of action for Rituximab have been confirmed including complement mediated cytotoxicity (CMC), antibody dependent cellular cytotoxicity (ADCC), cell cycle arrest and direct apoptosis induction.

CMC

The role of rituximab in CMC has been documented in several *in vitro* and clinical studies. In cell lines, the extent of *in vitro* lysis correlates with expression of complement inhibitory molecules (CD55 and CD59) on malignant B cells. In particular, CLL cells with greater expression of CD59 are resistant to CMC mediated by rituximab[157,158]. Furthermore, FL is clinically more sensitive to rituximab and these cells are more effectively lysed by complement *in vitro,* compared to large cell lymphoma or mantle cell lymphoma cells[159]. In animal models the importance of complement in anti-tumour activity was highlighted by depletion of complement effectors in animal models, diminishing the therapeutic response to Rituximab[159].

ADCC

Rituximab can induce ADCC which is mediated by several effector cells including NK cells, granulocytes and macrophages[160,161]. In this process the Fc region of Rituximab, which is bound to its target cell, binds to the Fcγ receptor on the effector cells initiating immune cell activation and death of target cell[162]. *In vitro*, rituximab was shown to induce ADCC of human lymphoma cell lines by human peripheral blood mononuclear cells (PBMCs)[163]. Studies of Rituximab have focused on its interaction with CD16 on NK cells although other Fc receptors likely involved include CD32, CD64 and CD89[164–166]. The strongest data linking ADCC to the clinical response of Rituximab comes from studies correlating polymorphisms of CD16 to clinical response to single agent Rituximab. CD16 homozygous for valine at position 158 has greater affinity for IgG1, and patients with FL

who have a VV genotype have a better clinical response to Rituximab than patients that have a VF or FF genotype[167–169].

Direct induction of apoptosis

Interestingly, Rituximab has also been shown to have significant non-immune related functions and has been shown to induce apoptosis via inhibition of several important signalling cascades. Studies have identified the p38 MAPK pathway as a potential target of Rituximab; Experiments in an AIDS related DLBCL cell line showed Rituximab selectively disrupted the IL-10 autocrine/paracrine loop via p38 MAPK signalling resulting in downregulation of IL-10, STAT3 and BCL2[170]. Other pathways that have been identified in modulating Rituximab induction of cell death include the Raf/MEK/ERK, NF-κB and the Akt signalling pathways. In BL cell lines, Rituximab selectively inhibited BCL-xL expression, leading to chemosensitization[171]. The molecular mechanism involved was identified as inhibition of the Raf1-MEK1/2-ERK1/2-AP-1 signalling pathway by Rituximab. As BCL-xL expression is under the transcriptional regulation of AP-1 it was consequently downregulated. Because NF-κB also regulates BCL-xL, the effect of Rituximab on this pathway was investigated in BL cell lines[172]. Decreased phosphorylation of NF-κB–inducing kinase (NIK), IκB kinase (IKK), and IκB-α was observed along with BCL-xL downregulation[170]. Finally it has also been reported that Rituximab inhibits the PI3K/Akt pathway in Ramos cells through inhibition of phospho-PI3K, PDK-1 and Akt. This inhibition was also shown to have downstream effects including inhibition of the NF-κB pathway and suppression of BCL-xL expression[170].

1.7.2.1 Resistance to Rituximab

Mechanisms of resistance to Rituximab *in vivo* remain poorly understood, although studies have reported several potential contributors include polymorphisms of the Fc receptor on cytotoxic cells, inhibition of CDC, downregulation or loss of CD20 expression on the surface of sub-clones, overexpression of anti-apoptotic proteins (e.g. BCL2), shedding of CD20-Rituximab complexes and failure to respond to Rituximab-mediated cell signaling[145,173–176]. Independent studies involving induction of acquired resistance to Rituximab *in vitro* showed that it is not associated with genetic changes in the CD20

molecule but both global gene and protein down-regulation of CD20 were noted[176–178]. In line with the preliminary evidence that several pathways may mediate the apoptotic effects of Rituximab, Bonavida *et al.* showed in their Rituximab-resistant clones that Rituximab failed to inhibit the p38 MAPK, ERK1/2, and NF-κB signalling pathways[170,176]. In addition, these resistant clones showed hyperactivation of these pathways, with overexpression of BCL2 and BCL-xL. Interestingly the cells also showed cross-resistance to other chemotherapeutics and inhibition of each of these pathways re-sensitized the resistant clones to various chemotherapeutic drugs including vincristine, etoposide and doxorubicin[179,176].

1.8 Intra-tumour heterogeneity

Intra-tumour heterogeneity (ITH) is one of the greatest challenges in cancer treatment leading to therapeutic failure, drug resistance and relapse[180]. ITH provides diverse genetic and epigenetic material, with varying levels of fitness, upon which selection and Darwinian evolution can act. The concept of genetic diversity within tumours dates back as far as 1958 to evolutionary biologist Julian Huxley who commented on "genetic inhomogeneity" in cancer[181]. However technical advancements such as next-generation sequencing were not in common use and it took many more years to decipher these observations. Early studies focused on the clonal nature of cancer; Evidence from Chronic Myeloid Leukaemia (CML) showed that each dividing cell carried the same chromosomal aberrations suggesting the malignancy had arisen from a single cell carrying this abnormality[182]. This model implies that a single cell acquires the 'first hit', in this case the BCR-ABL Philadelphia chromosome, possibly as a result of a random event or a carcinogenic insult[182,183]. The translated product is consistently expressed in all CML patients suggesting its causative role. However, evidence emerged that this alone was not sufficient for leukaemogenesis with as many as 30-70% of healthy individuals carrying the BCR-ABL mRNA transcript in their white blood cells[184–186]. It became obvious that additional mutations were necessary for full transformation and this random process gave rise to sub-populations of cells. The concept of sub-clonality was realised where cells may share a common clonal characteristic but can also individually acquire unique mutations. Even after malignant transformation, a cancer

remains dynamic and continues to evolve. This ongoing evolution might ultimately generate a molecularly heterogeneous bulk tumour consisting of cancer cells harbouring distinct molecular signatures with differential levels of fitness and sensitivity to anticancer therapies. This heterogeneity can result from genetic, transcriptomic, epigenetic, and/or phenotypic changes[180]. Numerous studies have highlighted the clinical implications of sub-clonal diversity particularly in the era of targeted therapies. In particular, work by Gerlinger and Swanton showed through multi-region sequencing and phylogenetic reconstruction, that 63 to 69% of all somatic mutations were not detectable across every tumour region in metastatic renal cancer, presenting significant challenges to personalized medicine strategies that depend on results from single tumour-biopsy samples[187].

Beyond the genetic level, phenotypic and functionally heterogenous populations in tumours have been well documented[188–191]. Different methods have been proposed to explain heterogeneity among cancer cells. The first model involves intrinsic differences caused by stochastic genetic or epigenetic changes[192,193]. The second model describes how cancers follow a stem cell model in which tumourigenic cancer stem cells "differentiate" into non-tumourigenic cancer cells, creating a hierarchical organization.

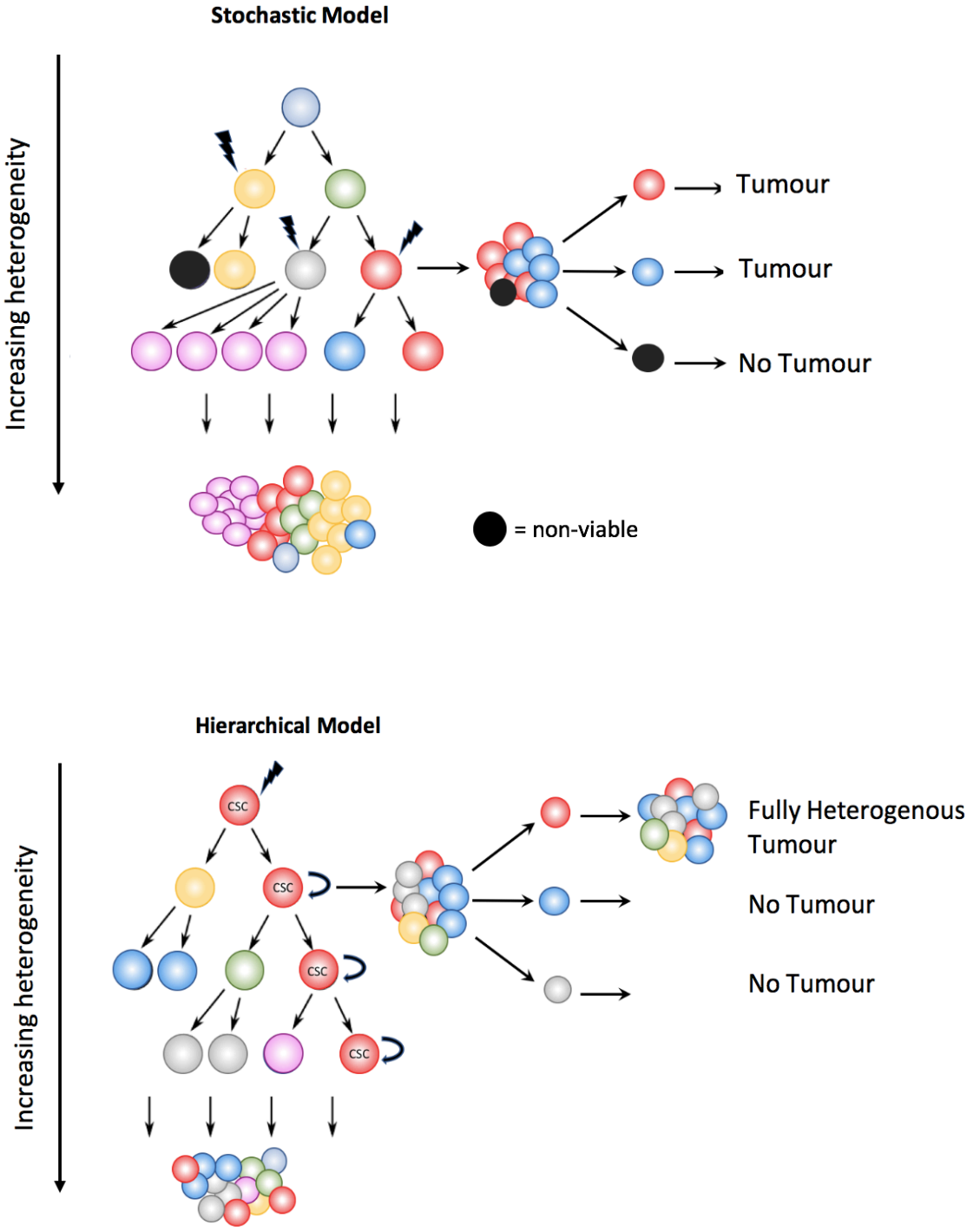


Figure 1.4. Schematic of the classical Stochastic and Hierarchical Models of tumour heterogeneity

1.8.1 Sources of intra-tumour heterogeneity

1.8.1.1 Stochastic model of tumour evolution

The field of cancer biology has been dominated by the stochastic model of tumour evolution. This model hypothesises that each tumour cell has equivalent replicative capacity and that mutations conferring a proliferative or survival advantage result in that particular cell and its progeny becoming the dominant clone in the tumour[192]. According to this theory almost every cancer cell could reproduce the whole cancer. Nowell formulated this hypothesis based on Darwinian principles describing cancer as an evolutionary process driven by stepwise somatic cell mutations with sequential selection of sub-clones, all of which are derived from a common progenitor. All surviving cells in the population have tumourigenic potential and hence a therapeutic regimen must target all cell populations in the tumour[191,192,194].

The very essence of multi-cellular and long-lived animals dictates the requirement for properties that, if not tightly regulated result in malignancy, for example, cellular self-renewal and stabilisation of telomeres[195]. In the majority of cases, the evolutionary progression of cancer is halted as highlighted by the high frequency of asymptomatic pre-malignant lesions[196]. Even advanced malignancies demonstrate Gompertzian growth indicating that the vast majority of cancer cells either die before they are able to divide, or are restricted from dividing by the tumour microenvironment. As is the case in natural selection within species, selection in tumours is driven by competition for space and resources. Individual clones evolve via the interaction of selectively advantageous “driver” lesions, selectively neutral “passenger” lesions and deleterious lesions. Further layers of complexity are introduced by “mutator” lesions which increase the rate of other genetic changes and by micro-environmental changes which alter the fitness effects of those lesions. Furthermore, the rate of epigenetic change is thought to be orders of magnitude greater than that of genetic change and is likely to be a major contributor to clonal evolution. Natural selection influences epigenetic variation within neoplasms because epigenetic changes are inherited at cell division and affect phenotype. Evidence to support the clonal evolution model is seen in cytogenic studies where all cells in many primary tumours show the same abnormal karyotype, implying

unicellular origin and even when several chromosome patterns are present within a single tumour, marker chromosomes in each cell often indicate that the various sub-populations derive from a common progenitor[191]. Despite the compelling evidence in favour of this theory, this model alone fails to sufficiently explain the cellular diversity comprising many tumours. As techniques to study functional tumour cell heterogeneity have improved, work involving patient-derived xenograft (PDX) tumour models highlight that functional tumour cell heterogeneity accompanies the phenotypic diversity observed within tumours. This work led to the emergence of the Cancer Stem cell paradigm for tumourigenesis in both haematological and solid tumour malignancies[197,198].

1.8.1.2 Cancer stem cell model

The cancer stem cell (CSC) model dictates that tumours are formed of a cellular hierarchy demonstrating some similarity to normal tissue, with a self-renewing, multipotent cell at the apex of the hierarchy that no longer responds appropriately to environmental cues[197,199]. This distinct set of stem cells undergo a slow, asymmetrical division to produce an identical stem cell while simultaneously producing a more differentiated progenitor cell. The progenitors then undergo rapid divisions to generate the bulk of the tumour[194,200]. The idea that tumourigenesis may be analogous to embryological development goes back to the 1800s, although it was the culmination of many years of research and the development of both xenografting techniques and Fluorescent Activated Cell Sorting (FACS) that led to the successful characterisation of CSCs in AML by Dick and colleagues[197,201]. Using FACS analysis of a variety of surface markers, they determined that not every AML cell was equal and that most significantly only rare cells were capable of tumour initiation. They referred to these cells as tumour or leukaemia initiation cells (LICs), the operational term for CSCs; AML is organised into a hierarchy with $CD34^{+}CD38^{-}$ cells at the apex. AML supports the hierarchical CSC hypothesis as it follows three key characteristics[202]:

1. Only a minority of cancer cells is endowed with tumourigenic potential when transplanted into immunodeficient mice,

2. Tumourigenic cancer cells are characterized by a distinctive profile of surface markers and can be differentially and reproducibly isolated from non-tumourigenic ones by flow cytometry, and
3. Tumours grown from tumourigenic cells contain mixed populations of both tumourigenic and non-tumourigenic cancer cells, thus re-creating the full phenotypic heterogeneity of the parent.

Clarke *et al*. added further detail to this definition of CSCs by stating that self-renewal must be demonstrated by serial passage in a xenograft assay at clonal cell doses and that the non-tumourigenic daughter cells produced may possess proliferative capacity but are unable to establish or maintain the tumour clone upon serial passage[200]. The term CSC is reserved for cases where a specific surface marker phenotype can be assigned to these cells[202]. If this is not possible, but the cells meet all other requirements of this definition, then they are termed TICs. These seminal studies in AML laid the ground work for the identification of CSCs in solid tumours 15 years ago with $CD44^{+}CD24^{-}$ cells in breast cancer shown to be responsible for sustaining the disease in immunodeficient mice[198]. Despite the fact that a plethora of studies followed supporting the CSC hierarchical model of carcinogenesis, the concept is still the subject of significant debate and has not been proven or disproven for many cancers including lymphomas. In particular, one of the key debated topics in the field of lymphoid malignancies is whether a series of gene alterations occurs in a healthy haematopoietic stem cell, or alternatively, a highly committed progenitor de-differentiates into a stem-like tumour-initiating cell[203].

1.9 Unified model of heterogeneity

Recent evidence has highlighted that the stochastic and hierarchical models are not mutually exclusive and it has been suggested that CSCs may not have to strictly fit the hierarchical model[204]. A more unified or harmonised model was proposed by John Dick *et al*. which centres on the concept of 'stemness'[202]. Stemness is defined as the functioning of molecular programmes that govern and maintain the stem cell state. Rather than CSC being a static well-defined cell type, cancer genetics, epigenetics and the microenvironment are coming together to influence stemness. In this reversible, cellular plasticity model, the stochastic and CSC model are unified by postulating that cancer cells can interconvert between stem cell and differentiated states known as directional interconvertibility[205]. Intrinsic tumour cell processes and/or stimuli within the tumour microenvironment could influence differentiated tumour cells to reacquire stem cell characteristics[206–208]. Conversely, these processes could also drive CSCs toward differentiation into non-stem cancer cells[207].

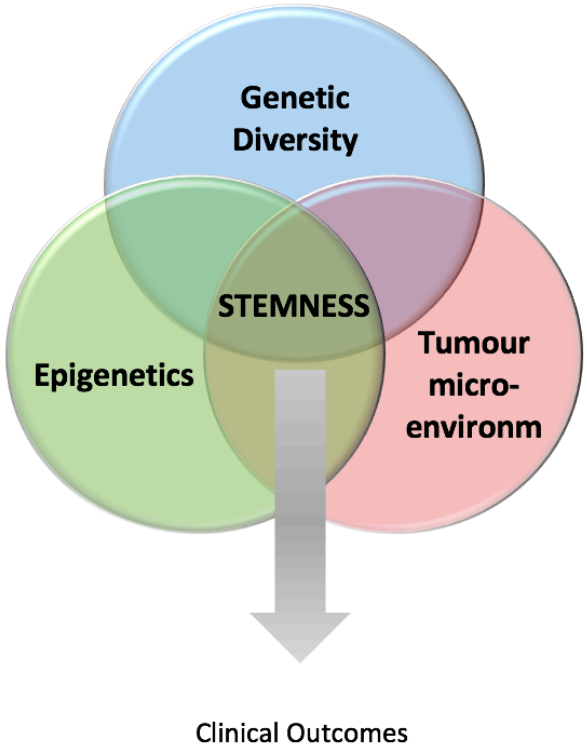


Figure 1.5. Three fields of cancer biology converge to influence stemness[202]

1.9.1 Defining stemness

The term stemness likely evolved as a result of the difficulty of assigning all characteristics of the CSC model to a population. In a review of the literature, Kreso and Dick consider the stemness state to consist of a variety of functions that enable cells to survive therapy regardless of whether the tumour is organised into a rigid hierarchy or not[202]. The determinants of stemness are the core contributors that affect therapy failure, irrespective of whether these determinants are present within a transitory state or in well-defined CSC populations. Genetics, epigenetics and the microenvironment all drive the stemness phenotype and the focus is on the molecular programmes that govern and maintain the stem cell state rather than a distinct cell phenotype. Despite these explanation considerable disparity still exists in the CSC field[202]. A recent book by Laplane made significant strides in combining all the relevant experimental literature — historical and scientific — as well as interviews with some of the leading international stem-cell researchers and clinicians[209]. She provided a framework of four possible versions of 'stemness': two intrinsic, two extrinsic. Stemness can be categorical (an intrinsic property of a stem cell, independent of its environment); dispositional (an intrinsic property of a stem cell that emerges only in the right environment); relational (an extrinsic property induced in a cell that would otherwise be a non-stem cell by its microenvironment); or systemic (an extrinsic property of a system such as tissue, rather than an individual cell). The requirement for such a framework highlights the complexity of this field and the significant work still required to both understand and clinically exploit the stemness population within different cancers.

1.9.2 CSCs and plasticity

Recent studies have provided evidence for the concept of cancer stem cell plasticity in which CSC and non-CSCs may not be definitive, but rather have a transitionary capability of shifting from a non-CSC state to a CSC state and vice versa[210–215]. *In vitro,* a number of TICs from cell lines were shown to be non-static as TIC-enriched populations generated non-TICs, and most significantly some studies found that sorted non-TIC populations could generate TICs[199,214,216]. In experimental models of melanoma,

tumourigenic cells display considerable plasticity, transiently acquiring stemness properties depending on the tumour context. In particular, both $CD133^+$ and $CD133^-$ melanoma cells have the ability to form tumours, suggesting that CD133 is reversibly expressed by tumourigenic melanoma cells rather than identifying cells at a static level in a hierarchy[217]. Further evidence of the complexity of TICs was also described in melanoma; JARID1B a histone demethylase, was shown to be essential for continuous tumour growth of established melanomas and metastatic progression[213]. JARID1B expression was limited to a small sub-population of melanoma cells, but experiments showed that the maintenance of this sub-population was dynamic: while purified JARID1B-positive cells generated JARID1B-negative cells, single JARID1B-negative cells also gave rise to heterogeneous progeny, including JARID1B-positive cells. This study indicates that some cells that are essential for tumour maintenance may not be static entities, but rather can acquire tumour maintenance capabilities depending on the context. This plasticity is driven by environmental cues that induce reprogramming or signalling events that drive stemness. For example, in mouse models of intestinal tumour initiation, epithelial non-stem cells can re-express stem cell markers upon Wnt activation driven by NF-κB, and are capable of ''de-differentiating'' to TICs[218]. A study by Weinberg and colleagues further supported the bi-directional nature of the CSC model. They showed the existence of sub-populations of non-CSCs that could readily switch to an aggressive CSC state[211]. This plasticity was induced by the transcription factor *ZEB*1 (a well-characterized EMT transcription factor) and microenvironmental stimuli could enhance the rate of non-CSC to CSC conversion[210]. In order to try to determine the reprogramming factors that may be responsible for dynamic stemness, Suva *et al.* introduced developmental transcription factors in a step-wise manner into differentiated glioblastoma cells. Remarkably, they showed that a combination of POU3F2, SOX2, SALL2 and OLIG2 resulted in cells capable of tumour initiation in 100% of animals by binding and activating CSC-specific regulatory elements resulting in the recapitulation of the epigenetic and phenotypic properties of CSC[219].

These insights and the potential impact of CSC plasticity towards the treatment of cancer dictate further study should be carried out across all cancers including lymphomas. B cells in particular inherently display extraordinary plasticity. For example,

differentiated B cells can be reprogrammed into macrophages whereby introduction of CEBPα/β expression inhibited the B-cell identity gene, *Pax5*[220]. Furthermore, studies of induced pluripotent stem cells (iPS) showed that terminally differentiated B cells could be reprogrammed to pluripotency through ectopic expression of Oct4, Sox2, Klf4 and c-MYC in conjunction with interruption of the B-specific transcriptional program by specific knockdown of Pax5 or exogenous expression of CEBPα/β[221]. These data suggest that B cells may be excellent substrates for reprogramming events causing induction of the CSC phenotype.

1.9.3 CSCs and the microenvironment

Cancers comprise malignant cells together with inflammatory cells, hematopoietic cells, associated stroma and vasculature. In line with the evidence of induced plasticity in the CSC phenotype, the microenvironment has been identified as playing an important role in providing the signals required for inducing stemness and maintaining the balance between CSC and non-CSC. The interaction between CSCs and the tumour microenvironment (TME) is often bidirectional; CSCs in glioblastoma have been demonstrated to secrete VEGF that modifies the microenvironment and directly supports the development of the local vasculature[222]. In the opposite direction endothelial cells secrete nitric oxide that induces Notch signalling in glioma cells and drives stemness[223]. In squamous cell carcinomas, CSCs in the vascular niche establish an autocrine loop in which VEGF promotes CSC activity by governing both the microenvironment and intrinsic self-renewal pathways in CSCs[224]. Evidence of the direct effect of immune cells on CSCs was provided in a study of colorectal cancer; $CD4^+$ T cells were shown to produce IL-22 which acted through STAT3 to induce the stem cell genes NANOG, SOX2, and POU5F1. This resulted in increased cancer stemness and tumourigenic potential[225]. Other cytokines that have been shown to have an impact on cancer cell stemness and plasticity include TNF and IL-6[226]. In melanoma, breast and lung cancer studies these pro-inflammatory mediators impact the differentiation state of tumour cells by up-regulating genes involved in EMT and in breast cancer, activation of production of IL-6 induces malignant features in NOTCH3–expressing stem/progenitor cells[227,228].

Most studies have focused on the interactions between CSCs and host stromal and immune cells, although recent studies have begun to shed light on the communication between CSC and non-CSCs. Zhang *et al*. identified a positive feedback loop between $CD29^{high}CD24^{high}$ CSCs in breast cancer and mesenchymal populations derived from them[229]. The non-CSCs produce signalling molecules including Wnt2, Wnt9a, CXCL12, and IL-6 - these molecules then promote self-renewal of the CSCs. Since the $CD29^{high}CD24^{low}$ mesenchymal cells were derived from the $CD29^{high}CD24^{high}$ CSC population, this data highlights that CSCs are able to generate their own niche. Further evidence of communication between CSCs and non-CSCs came from a study of DLBCL[230]. Koch *et al.* isolated a putative CSC populating using the side population assay. They demonstrated that CSC and non-CSC populations could transition from CSC to non-CSC as well as from non-CSC to CSC and that the conversion of cells states was modulated by exosome-mediated Wnt signalling specifically from the CSC population.

Similar to the discussions on plasticity, the potential for B-cell lymphoma cells to be impacted by their microenvironment is significant. GC and memory B cells, the cells of origin of most lymphomas, reside in GCs where they interact with stromal cells and differentiate into mature cells that possess self-renewal potential that allows rapid lymphocyte division and antigen-specific immunity. Interestingly, as well as sharing self-renewal capabilities with haematopoietic stem cells (HSCs), memory B cells share a transcriptional program of self-renewal with long-term HSCs including expression of BMI-1, a stemness related gene in numerous cancers[231]. The renewal and differentiation capabilities of GC B cells upon antigen stimulation make secondary lymphoid organs ideal environments to confer stemness through stroma-mediated genetic and epigenetic mechanisms. The close physical contact and interactions of lymphoma cells and stromal cells in the lymph node and bone marrow microenvironment provides an appropriate niche for malignant reprogramming to be achieved through stromal-mediated regulation of genes controlling stemness pathways. Indeed, studies have shown that stromal cells can promote stemness in B-NHL cells by influencing B cell differentiation and clonogenicity[232,233]. Furthermore, stromal cells can also modulate key stemness genes including *c-Myc, EZH2, Oct-4, Nanog* and *Sox2*[234,235].

1.9.4 Pathways regulating stemness

In both the classical unidirectional CSC model and the plastic bidirectional CSC model, a number of pathways have been identified as driving the stemness phenotype in CSCs. This list is constantly growing and current evidence supports roles for the key developmental pathways Wnt, Hedgehog and Notch in mediating key stemness properties. Other implicated pathways include JAK/STAT, Hedgehog, Wnt, Notch, PI3K/PTEN, NF-κB and MAPK pathways[236,237].

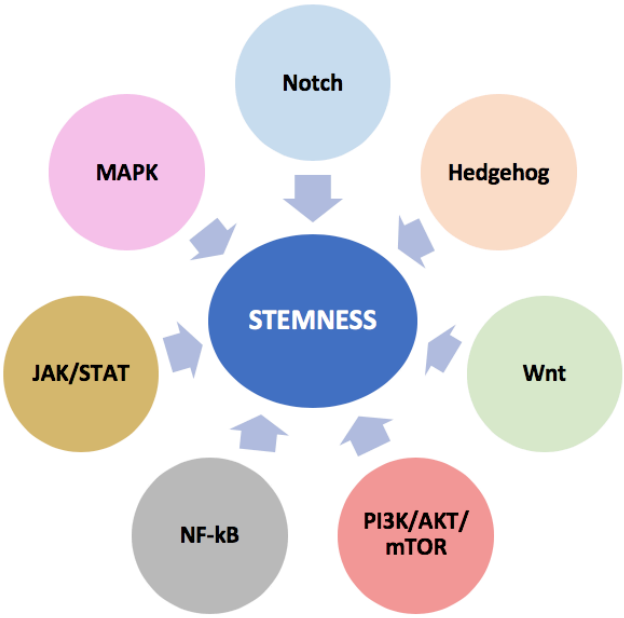


Figure 1.6. Numerous different pathways have been implicated in promoting stemness

1.9.4.1 Wnt signalling

The Wnt pathway is a vital, evolutionarily conserved pathway in embryonic development and tissue homeostasis, in which Wnt ligands signal through β-catenin for their biological functions via binding to the seven-transmembrane receptor, Frizzled (FZD) and the single-membrane-spanning low-density receptor-related protein 5/6 (LRP5/6)[238]. Wnt has been shown to drive stemness in a range of cancers including colorectal, breast, hematologic, skin and lung cancers[239–243]. In particular, Wnt signalling

has been shown to be a key driver of CSCs in colorectal cancer. Abnormal activation of Wnt signalling disrupts the normal growth and differentiation of colon crypt stem cells, resulting in a colorectal CSC phenotype by upregulating expression of target genes such as MYC and cyclin D[244,245]. In squamous cell carcinoma, Wnt signalling activation was shown to be critical in tumourigenesis of $CD34^+$ CSCs, and knock down of the β-catenin gene resulted in depletion of $CD34^+$ CSCs and complete regression in mice[243]. LEF1, a component of the Wnt signalling pathway was recently detected at different levels in a large portion of DLBCL samples and it was also identified as a signature gene and possible therapeutic target in the molecular BL subtype [246,247].

1.9.4.2 Hedgehog signalling

The Hedgehog (HH) pathway has also been shown to be important in the maintenance of CSCs in an array of cancers including glioblastoma, lung squamous cell carcinoma, breast cancer, pancreatic adenocarcinoma, myeloma, and CML[248–251]. HH signalling plays an important role in plasticity and de-differentiation during the EMT process to acquire stem cell-like phenotypes, and inhibition of HH signalling by SMO knockdown in pancreatic cancer blocks the self-renewal, EMT, invasion, chemoresistance and tumourigenesis of CSCs[252]. HH signalling components have also been shown to drive stemness in lymphoid malignancies. In multiple myeloma, SMO and GLI1 were highly expressed in the CSC compartment compared to non-CSCs; Activation of HH signalling by ligands induced CSC expansion, while inhibition of HH signalling significantly blocked CSC expansion[253]. In a murine model of CML, deletion of SMO significantly reduced CSCs, and overexpression of SMO in a SMO-deficient mouse CML model resulted in a 4-fold enhancement of CSCs with simultaneous disease progression[251]. Interestingly, in BL MYC was shown to be able to regulate GLI activity independent of HH ligands or the SMO and PTCH signalling molecules[254].

1.9.4.3 Notch signalling

The Notch pathway is another key evolutionarily conserved pathway in development and most significantly it regulates the expression of target genes including some that are essential to CSC self-renewal such as *MYC, NANOG, OCT-4*, and *SOX2*[255–258]. Notch

signalling has been shown to be instrumental in promoting CSCs of breast cancer, pancreatic cancer, and glioblastoma; In a study of 115 breast tumour tissues from primary lesions, results showed that Notch-positive tissues were significantly associated with the CSC marker ALDH1[259]. In a study of CSCs from pancreatic patients, Notch ligands Notch 1, Notch 3, Jag1, Jag2, and Notch target gene *Hes1* were found to be highly expressed, and inhibition of γ-secretase (an important protease mediating Notch signalling by releasing the Notch intracellular domain) significantly decreased the CSC population and sphere formation abilities[260]. Finally, in Glioblastoma patient CSCs, Protein Kinase C Iota (PKCi) was shown to be highly expressed and important for CSC survival mediated by Notch signalling, and silencing PKCi caused apoptosis and inhibited proliferation of the CSCs both *in vitro* and in an *in vivo* PDX model[261].

1.9.4.4 JAK/STAT signalling

Modulation of the JAK/STAT pathway has been implicated in the enhancement and repression of CSCs in several cancers. In AML, analysis of isolated CSCs from patients showed constitutive activation of JAK/STAT signalling. *In vitro* studies highlighted that the growth and survival of these CSCs were inhibited following treatment with a JAK1/2 inhibitor[262]. Significant evidence for the impact of this pathway on glioblastoma CSCs has also been reported. TGF-β was shown to regulate the self-renewal and differentiation properties of glioma-initiating cells from patient samples and induced activation of the JAK/STAT pathway via leukaemia inhibitory factor (LIF)[263]. Furthermore, inhibition of STAT3 in CSCs prevented proliferation and sphere formation, and decreased expression of the neural stem cell genes *Olig2* and *nestin,* while simultaneously increasing expression of the neuronal differentiation gene *βIII-tubulin*[264]. Indeed STAT3 has been strongly implicated in promoting the stemness phenotype in an array of cancers. In breast cancer, it mediates resistance of $CD44^{+}CD24^{-}/^{low}$ breast CSCs to tamoxifen *in vitro* and also increases the CSC pool in HER2 overexpressing breast cancer[265,266]. It was also reported that hepatocellular carcinoma could arise from STAT3-driven transformed stem cells with inactivated TGF-β signalling[267]. Importantly, IL-6 has been shown to be important in modulating STAT3-driven stemness. For example, in breast cancer, activation of an IL-6 inflammatory loop

mediates trastuzumab resistance by expanding the CSC population via upregulation of STAT3 (as well as AKT and NF-kB)[268]. IL-11, a less-well studied member of the IL-6 family has also been shown to influence breast CSC dynamics through STAT3[269].

1.9.4.5 PI3K/Akt/mTOR signalling

A wealth of evidence supports the role of the PI3K/Akt/mTOR pathway in modulating stemness. Inactivating mutations in PTEN are found commonly in glioblastoma multiforme and a recent analysis highlighted the initiating role of PTEN in the transformation of neural stem cells into glioblastoma multiforme while retaining the ability to express key stemness regulators OCT4, SOX2, and NANOG, and to differentiate into multiple lineages[270]. Another study in AML established a role for the microRNA miR-126 in the maintenance of stemness through PI3K signalling[271]. In breast cancer stem-like cells activation of the mTOR pathway is necessary for colony-formation ability *in vitro* and tumourigenicity *in vivo* while inhibition of mTOR in colorectal cancer caused a decrease in ALDH1 activity, a marker for colorectal CSCs[272–274]. In both ALL and AML, activated PI3K upregulated ABCG2 expression and elevated the percentage of cancer stem-like cells[275]. Bleau also presented similar results in Glioma stem-like cells and further elucidated that Akt but not its downstream target mTOR regulated ABCG2 expression[276]. This pathway has also been shown to regulate the expression of important signalling molecules related to stemness. In studies of NSCLC, abnormal activation of the PI3K/Akt/mTOR signalling pathway leads to enhanced expression of chemokine (C-X-C motif) receptor 4 (CXCR4), which in turn promotes CXCR4-mediated STAT3 signalling that may be responsible for maintenance of stemness[277]. Chang *et al*. determined that PI3K/Akt/mTOR signalling via the insulin-like growth factor-1 receptor (IGF-1R) are attractive targets for therapy directed against breast CSCs[278].

1.9.4.6 NF-κB signalling

The role of NF-κB in self-renewal and multipotency is less clear although evidence has started to emerge. For example, in a breast cancer pre-clinical study using a mouse mammary cell line and a mouse mammary model harbouring active ErbB2, investigators showed that inhibition of NF-κB reduced the number of progenitor cells and the ability

of these cells to form 3D mammospheres in culture. Furthermore they suggest that NF-κB is regulating this mammary -initiating cell population via SOX2 and NANOG as chemical inhibition of NF-κB decreased the promoter activity levels of these genes by approximately 90% and 50%, respectively[279]. NF-κB has also been reported to enable metastasis of CSCs; A recent study demonstrated that breast CSCs may promote metastasis to lungs via LIN28, a stemness factor and downstream effector of IKKβ[280]. *In vitro* analysis of breast cancer cell lines highlighted that breast CSCs expressed higher protein levels of LIN28 and the activated form of IKKβ than non-CSCs. Genetic silencing or chemical inhibition of IKKβ reduced expression of the stemness proteins LIN28, OCT4, SOX2 and NANOG. Finally, treatment of breast cancer xenograft-bearing mice with an IKKβ inhibitor almost completely eliminated the incidence of lung metastasis, and significantly reduced the frequency of CSCs in metastatic foci, compared with vehicle control treated cells[280]. TNF-α was also shown to trigger proliferation of adult neural stem cells via IKK/NF-κB signalling[281].

1.10 Cancer stem cells in lymphoma

The original CSC concept has been fully validated in AML where the initial oncogenic hit occurs in a healthy, tissue-specific stem cell[282]. However, whether CSCs exist in mature B-NHL has yet to be determined. Akin to models of leukaemogenesis, potential B-NHL stem cells could derive either from HSCs or multipotent progenitors, or alternatively reprogramming events induced by cell intrinsic or environmental factors could induce stemness in mature B cells.

In favour of the former model, several case studies report apparent transmission of premalignant lymphoid cells from one individual to another. At least 7 cases have been reported in which both recipient and donor develop the same lymphoid disease after allogeneic transplant[283–288]. Importantly, each case also demonstrates genetic sequence homology between the tumours of the recipient and the donor, thereby excluding a chance finding. The cases include T-cell Lymphoma (TCL), Marginal Zone Lymphoma (MZL), CLL, FL and Mantle Cell Lymphoma (MCL) although no cases of DLBCL or BL have been reported with these observations. In one case of MCL, onset of disease occurred

simultaneously in both donor and recipient 12 years after transplant for AML[283]. The malignant clone in the female recipient was found to be derived from the male donor (tumour with Y chromosome) supporting the existence of a latent premalignant clone and also providing evidence for the important role the TME has in promoting or repressing transformation. In one case of FL, a different IgH sequence was found in the 2 tumours supporting the hypothesis of a common lymphoma progenitor cell existing[285]. Determining what factors promote the development of a putative lymphoma CSC into active proliferative cancer remain to be elucidated.

Sequencing studies in healthy individuals have shed further light on the concept of premalignant lymphoma CSC clones. Two seminal papers found that specific somatic mutations in peripheral blood cells are detected in healthy adults and precede the onset of hematologic malignancy[289,290]. The frequency increases by age; by age 70, ~10% of healthy individuals have common aberrations. Some of the common mutations involve epigenetic regulator genes which are more common in myeloid cancers, but also occur in lymphoid malignancies[291–293]. In CLL it has been shown that over ~90% of patients harbour the same mutations in their $CD34^+$ stem cells as in their malignant $CD19^+$ bone marrow cells[294]. Furthermore, in Hairy Cell Leukaemia, the BRAF V600E mutation which is found in the majority of patients was also discovered in their $CD34^+$ stem cells[295].

Despite these fascinating discoveries no similar results have been reported for BL and the data for DLBCL are less conclusive. A recent study of 60 DLBCL patients identified copy-number neutral loss of heterozygosity in key lymphoma regions in non-tumour control samples[296]. These regions included 6p21 (HLA, 3% of patients), 9p24.1 (PD-L1/L2, 5% of patients), and 17p13.1 (*TP53*, 2.5% of patients). However, in another study deep sequencing of genes frequently mutated in DLBCL revealed no mutations in highly purified HSCs from 6 DLBCL patients[297]. Furthermore, sequential sequencing was carried out on presentation and relapse samples from these patients and although common somatic mutations and clonal IgH VDJ rearrangements were revealed, none were identified in the purified patient HSCs. This may suggest that the source of DLBCL relapses is of a more mature B-cell type than in low-grade B-cell malignancies. In a detailed study of the methylation profile of 83 mature aggressive B-NHL samples,

including BL, investigators identified 56 genes *de novo* methylated in all lymphoma subtypes studied, and 22 methylated in a lymphoma subtype-specific manner[298]. Interestingly, the group of genes *de novo* methylated in all lymphoma subtypes was significantly enriched for polycomb targets in embryonic stem cells. *De novo* methylated genes in all cases studied were expressed at low levels in lymphomas and normal hematopoietic tissues, but not in non-hematopoietic tissues. These findings, especially the enrichment for polycomb target genes in stem cells, indicate that molecular B-NHLs with different morphologic, genetic and transcriptional background share a similar stem cell-like epigenetic pattern. This provides evidence for two alternative explanations for stemness in mature B-NHLs: On one hand these diverse B-NHLs might derive from lymphoma precursor cells with similar stem cell-like features that have acquired aberrant methylation of PcG target genes, or alternatively an epigenetic footprint of stemness could be acquired during lymphomagenesis by epigenetic remodelling with common patterns across different lymphomas. Within the molecular reprogramming model of stemness in mature B-NHL it is possible that epigenetically aberrant lymphoma initiating stem or progenitor cells may retain or reacquire stem cell features. They are then susceptible to become fully transformed via the acquisition of translocations such as MYC, BCL2 and BCL6. Alternatively the translocation breakpoint product or other genetic aberrations may constitute the first hit and subsequent reprogramming events confer stemness.

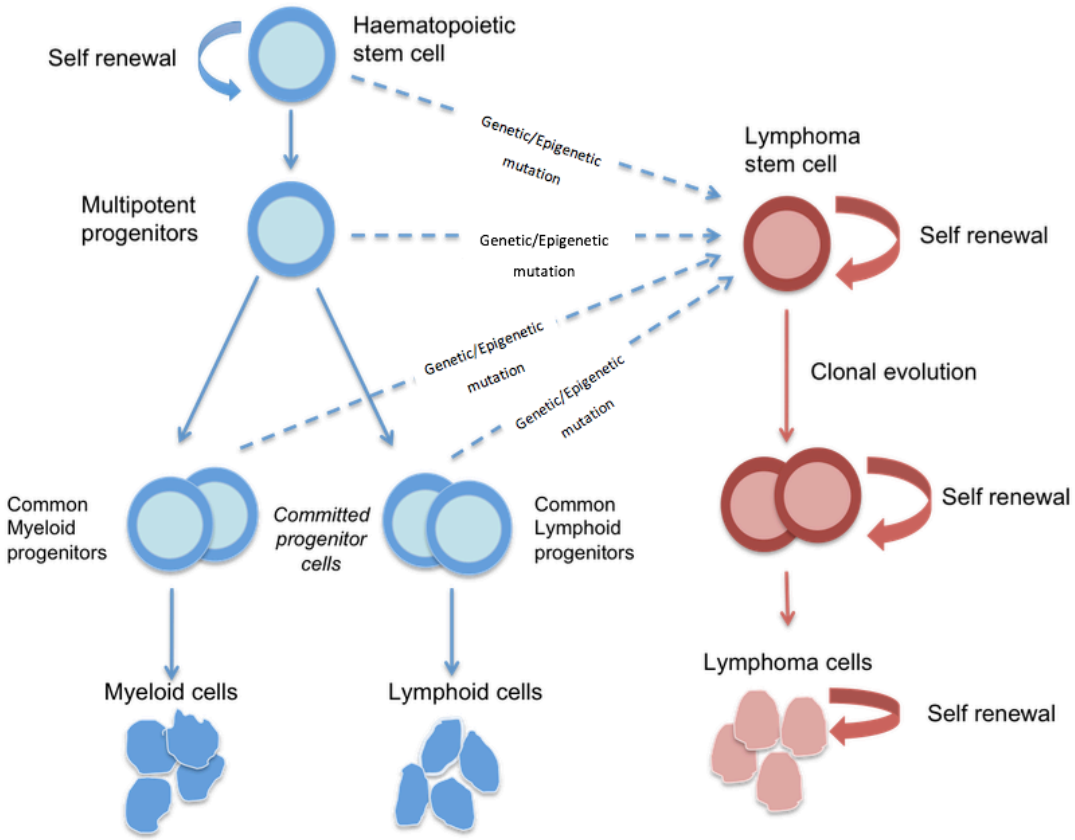


Figure 1.7. Potential models of stemness in Lymphoma

Some Lymphomas show evidence of a potential initial oncogenic hit in a HSC while for others such as BL and DLBCL it is more likely that genetic and/or epigenetic perturbations and the microenvironment confer stemness on more mature lymphoid cells. Adapted from [203].

1.11 Methods of Identifying CSCs

1.11.1 Xenotransplantation assays

The gold standard method for identifying TICs is the use of *in vivo* tumour propagating assays. Suspected CSCs are assessed for their ability to serially propagate fully heterogenous tumours compared to the CSC-depleted bulk tumour population. Furthermore, this assay allows for the transplantation of different numbers of proposed CSCs into immunodeficient mice, and the subsequent establishment of a linear correlation with engraftment can determine a quantitative estimate for the frequency of CSC in the tumour. The first studies by Dick and colleagues determined that 1 in 2.5 X 10^5 cells could initiate a leukemic graft[197]. Because xenografting is such an important method involved in defining CSCs it has come under intense examination over the years and some of the limitations have been described. One of the most obvious limitations is the inability of transplantation methods to faithfully recapitulate the TME. This is particularly relevant in the modern era of CSC studies whereby a cell's environment is known to play an important role in dictating stemness features. Many murine growth factors are not cross-species reactive (e.g. TNF) and the possibility that some genuine TICs in humans may lack the appropriate conditions for growth in animals must be considered[299,300]. Furthermore, the experimental techniques utilised to obtain single cells for TIC assays are harsh and can have negative implications for cells. Both chemical and mechanical digestion methods are common, during which cells are subject to atmospheric 0_2 as well as sub-optimal pH and nutrient levels. They are then usually stained with antibodies and subjected to cell sorting via FACS. The significantly different environment that cells encounter for several hours before injection into mice may apply selective pressure to the population favouring only the most robust cells and may also induce epigenetic changes.

Despite these complexities over the past several years a number of improvements have been achieved in xenotransplantation including development of more immune-deficient recipient mice, better methods for transplantation, introduction of basement membranes and humanizing recipients with human TME and/or growth factors[300,301].

Many initial discoveries were made using NOD/SCID (NOD.CB17-Prkdc^{SCID}/J) mice. These mice have defective cellular and humoral immunity but the presence of functional NK cells has been shown to impact the efficiency of the systems[302,303]. This was particularly evident in studies by Quintana *et al.* whereby the NOD/SCID model was shown to underestimate the CSC frequency[304]. When more immune-deficient NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ (NSG) mice were used for xenotransplantation in melanoma, the frequency of CSC cells was estimated at 25% compared to the 0.1-0.0001% in the NOD/SCID model. The main difference in NSG mice is that the introduction of truncated IL2Ry hampers the production of NK cells[303,305]. This model has become the most commonly used in xenotransplantation studies and has shown increased engraftment rates across many cancer types.

The introduction of the basement membrane Matrigel has also had a significant impact on efficiency of engraftment in xenotransplantation assays. Matrigel is a solubilized tissue basement membrane matrix rich in extracellular matrix proteins that was originally isolated from the Engelbreth-Holm-Swarm (EHS) mouse[306]. In light of the important role the TME plays in modulating stemness, the ability of matrigel to apply a more tumour specific niche may be particularly important in CSC studies. It is composed mainly of laminin, collagen IV, heparan sulphate, proteoglycans and entactin (nidogen), although various growth factors such as transforming growth factor-β (TGF-β), fibroblast growth factor (FGF), and tissue plasminogen activator are also present which may play important roles in CSC progression[306]. Furthermore, at 37^{o}C matrigel forms a 3D gel providing a scaffold for tumour cells[307]. This may be particularly important in tumours where cell to cell physical interaction and communication are important. In melanoma, initial estimates of the frequency of CSCs was increased significantly when matrigel was used[304]. The long list of studies that have utilised basement membrane proteins in the identification of TICs is evidence of the important role this component has in supporting CSCs[304,308–316]. It is also striking how effective matrigel is at promoting tumour growth from very low numbers of cells. In fact, in human melanoma one TIC cell injected with matrigel successfully produced a fully heterogenous tumour[304]. Similar results were seen for human squamous cell carcinoma with TGF-B shown to be involved in controlling the TIC phenotype[308].

1.11.2 Cell surface phenotyping

The use of flow cytometry to identify putative CSCs has been a very powerful technique over the years, particularly due the ability to integrate FACS to separate cell compartments based on a combination of cell surface markers. Surface marker expression can be hugely informative regarding the differentiation state of cells within defined lineages. In the normal stem cell field, HSCs were shown to have a surface phenotype denoted originally by $CD34^+CD38^-$ and further extended to include Thy-1 (CD90), IL7Ra and Flk2[317]. It was knowledge of the $CD34^+CD38^-$ phenotype that contributed to the seminal studies by Dick *et al.* in identifying leukemic stem cells in AML[282]. However, as research in the area has developed some discrepancies with the original findings were identified. Initially LICs were thought to reside exclusively in the $CD34^+CD38^-$ fraction of AML[282]. However, following the use of NSG rather than NOD/SCID mice, LICs were identified in other cell fractions too[318]. Thorough investigations were carried out using larger sample numbers and the original conclusion was validated – in all cases the $CD34^+CD38^-$ fraction denoted a CSC population although, contrary to previous knowledge, in at least half of the more than 100 samples analysed, LICs were also found in other fractions[319]. In a study of 16 AML patients, 13 patients had tumour initiating capacity in fractions other than $CD34^+CD38^-$ ($CD34^+CD38^+$, $CD34^-CD38^+$ or $CD34^-CD38^-$). However, strong evidence for a hierarchical structure still remains as in the vast majority of cases, populations devoid of leukaemia initiating capacity exist. Furthermore, only gene signatures from functionally validated LIC populations were prognostic for patient survival[319].

Similar results regarding cell surface protein expression exist in other cancers. The original identification of the $CD44^+CD24^-$ phenotype as the CSC population in breast cancer has been contested and other populations have been identified that possess stemness characteristics including cells expressing ALDH, CD133 and CD49F[198,315,320,321]. A number of cell surface proteins in ovarian cancer have been reported as denoting CSCs including $CD44^+$, $CD24^+$, $CD133^+$, $CD117^+$ and $CD44^+CD117^+$[322–325]. However Stewart *et al.* report significant instability of the ovarian TICs and determined that $CD44^+CD117^+$ cells were absent in most ovarian cancer patients or were found not to be tumourigenic[326] .

These studies highlight some interesting situations – there appears to be significant interpatient heterogeneity with respect to CSC cell surface phenotype, a daunting concept in terms of targeting these populations for therapy. Furthermore, it appears that cell surface markers may be uninformative and may not represent a valid approach to determine hierarchical organization in some cases. Alternative methods of identifying and isolating cells with stemness may be required based on functional characteristics. Two potential options which have only recently been investigated include the use of sorting based on differential expression of miRNA, or the use of reporter assays that measure cellular signalling pathway activity[327]. Vermeulen *et al.* used this latter method in colon CSCs that were defined by Wnt signalling and regulated by the TME[328].

1.11.3 Side population assay

Another approach that focuses on functional characteristics to identify CSCs is the Side Population (SP) Assay which is based on the differential ability of cell populations to efflux Hoechst dye. The ability to expel this dye is mediated by the ATP-binding cassette (ABC) family of transporter proteins expressed within the cell membrane. It is estimated that between 500 to 1200 genes encode drug transporters but ABCB1 (P-glycoprotein, multidrug resistant protein 1, MDR1), ABCC1-5 (multidrug resistant-associated proteins, MRP1-5), and ABCG2 (breast cancer resistance protein, BCRP1) are the most well characterized transporter proteins that function to denote the SP phenotype[329–333]. The SP phenotype was first described by Margaret Goodell and colleagues in 1996 using mouse bone marrow cells[334]. The SP did not exclusively encompass the stem cell population although it overlapped with the phenotypically defined $CD117^{+}$ $Sca\text{-}1^{+}$ Lin^{-} $Thy1^{lo}$ HSC population, and most significantly when surface marker expression and the SP phenotype are combined, the resulting isolated cells are one of the most pure and potent adult stem cell populations available[335,336]. Since its discovery over 20 years ago, this powerful assay has been implemented in identifying stem cells and progenitor cells in an array of tissues including umbilical cord blood, skeletal muscle, kidney, liver, mammary glands and lung[337–342]. The field of cancer stemness then began to utilise the SP assay as a means of isolating this therapy-resistant population and indeed successfully identified SP cells have been implicated in chemotherapy resistance in a

large number of cancers[343–348]. Importantly, SP cells have also been shown to possess the other key characteristics of TICs – self-renewal and tumourigenicity in xenotransplantation assays in a plethora of malignancies[276,343,348,349]. Important considerations when employing the SP assay are that SP cells do not exist in every tumour. Reports have shown that dye efflux is not a defining property of all stem cell populations and hence if an SP phenotype is not present it does not mean that the tumour does not contain populations with stemness[350,351]. Furthermore, ABC transporter expression is not exclusive to the SP and indeed, as discussed above in relation to multidrug resistance, ABC transporters are expressed by specialized cells within several organs, including the small intestine, liver, kidney, brain and placenta[331,352–355]. Taken together it is important to emphasize that the term 'SP cell' and 'CSC' are not interchangeable. Rather, the SP phenotype offers a useful assay to enrich for cells with CSC potential and should be combined with cell surface phenotyping and xenotransplantation in order to further characterise the true CSC population.

Operationally, the SP is identified by staining cells with Hoechst dye and then visualising the Hoechst profile simultaneously in two distinct channels of a flow cytometer equipped with a UV laser. Hoechst 33342, when excited by UV light, emits fluorescence that can be detected in two distinct channels on the flow cytometer: the ''Hoechst Blue'' and the ''Hoechst Red'') channel, while a dichroic mirror is used to split the emission wavelengths. As the ''Hoechst Red'' channel is more sensitive to small changes in dye concentration, the so called ''side population'' cells emerge as a distinct dim ''tail'' extending first on the left side of G0/G1 cells (often referred to as the main population) toward the lower ''Hoechst Blue'' signal. This population is lost upon inhibition of ABC transporter activity with drugs such as verapamil and this offers a useful control for experiments[334,356].

Functionally, the SP phenotype is dynamic and is regulated by many growth factors, signalling pathways and micro-RNAs[357]. One example of this is seen in the haematopoietic system where the SP phenotype is acquired or reactivated in Lin- cells following interaction with mesenchymal stromal cells through VLA-4/α4β1 integrin and CD44[358]. Another study demonstrated that TGF-β reduced the SP in a dose-dependent

manner in hepatic cancer cell lines while Liu *et al.* reported that in NSCLC and colon cancer cell lines glucose induces a reversible upregulation of the SP fraction through ATP mediated suppression of AMPK and activation of the Akt pathway[359,360]. As well as this, ABCG2 transcription is regulated by several transcription factors that bind to the promoter region including Hif-1α, the oestrogen receptor, progesterone receptor isoform β, GLI, NRF2, SP1 and SP2[357]. Interestingly, several studies have also shown that a number of MicroRNAs negatively regulate the translation of the ABCG2 protein reducing the SP phenotype[361–363].

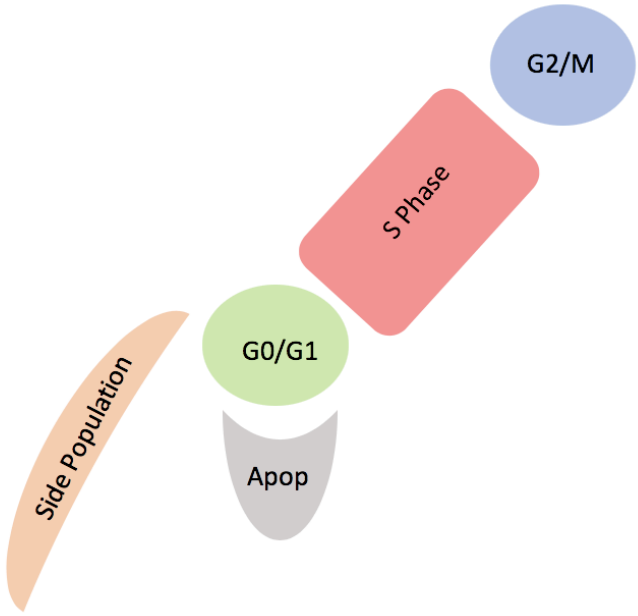


Figure 1.8. Identification of a side population by Hoechst staining

Following staining of cells with Hoechst 33342 dye the dye fluorescence signal can be detected in two distinct channels on the flow cytometer. As the ''Hoechst Red'' channel is more sensitive to small changes in dye concentration, the so called ''side population'' (SP) cells emerge as a distinct dim ''tail'' extending first on the left side of G0/G1 cells (known as the main population) toward the lower ''Hoechst Blue'' signal.

1.12 Aims of the investigation

The introduction to this thesis has given an overview of our current understanding of the biology of paediatric B-NHL as well as treatment options and the significant challenge relapse and refractory disease still poses. There is an unmet need for novel therapeutics to improve salvage rates of these patients as well as attempt to reduce the toxic doses of chemotherapy currently used. This thesis will investigate the potential importance of intra-tumour heterogeneity in paediatric B-NHL and in particular determine if a population with stemness properties exist which may contribute to therapy resistance. We will also utilise novel *in vitro* and *in vivo* methods to investigate resistance mechanisms and attempt to identify novel targets that may be effective in these difficult to treat cases.

Specifically, the aims of this thesis are to:

- Establish the first comprehensive paediatric B-NHL Patient Derived Xenograft (PDX) resource and determine if these offer a good model to investigate heterogeneity.
- Investigate both inter- and intra-tumour heterogeneity within the PDX resource and identify populations that may have therapeutic relevance.
- Determine if a population with stemness characteristics exists in paediatric B-NHL by xenotransplantation assays, cell surface phenotyping and functional assays and investigate the characteristics of this population both *in vivo* and *in vitro.*
- Investigate resistance mechanisms *in vivo* by attempting to mimic relapse in PDX models.
- Investigate resistance mechanisms *in vitro* utilising a CRISPR-Cas9 genome-wide over expression screen to provide rational targets for investigation in PDX models.

Chapter 2. Materials and Methods

2.1 Cell lines

Bjab, DG-75, Raji and Ramos cell lines were obtained from a validated collection at the DSMZ German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). Cells were grown in Roswell Park Memorial Institute (RPMI) 1640 (Gibco) with 10% Fetal Bovine Serum (FBS) (Labtech) and 1% Penicillin-Streptomycin (Invitrogen). All cells were grown in a humidified 5% CO_2 incubator at 37°C.

2.2 Primary Patient Samples

2.2.1 Patient sample acquisition

Paediatric B-NHL primary patient samples were obtained following informed consent on enrolment of patients to the Inter B-NHL Ritux Clinical Trial or following local ethical approval (granted to Dr. Suzanne Turner; Research Ethics Committee Reference: 07/q0104/16). Two samples were kindly provided by Dr. Simon Bomken from the Wolfson Childhood Cancer Research Centre, Northern Institute for Cancer Research, Newcastle University once they had been passaged through a mouse and hence not under the regulations of the Human Tissue Act.

2.2.2 Patient Sample Processing

Mononuclear cells (MNCs) were isolated from pleural effusion, peripheral blood or bone marrow by gradient centrifugation over Lymphoprep™ (STEMCELL Technologies). Depending on the volume of material received, the sample was mixed with PBS (containing 2% FBS) and carefully layered over Lymphoprep in volumes shown in Table 2.1.

BLOOD (mL)	PBS + 2% FBS (mL)	LYMPHOPREP™ (mL)	TUBE SIZE (mL)
1	1	1.5	5
2	2	3	14
3	3	3	14
4	4	4	14
5	5	10	50
10	10	15	50
15	15	15	50

Table 2.1. Lymphoprep and PBS volumes utilised for isolation of MNCs

Sample was collected by centrifugation at 800 x g for 20 minutes at room temperature (15 - 25°C) with the brake off. The upper plasma layer was removed and discarded without disturbing the plasma:Lymphoprep™ interface. The MNC layer was removed at the plasma:Lymphoprep™ interface, by pipetting, without disturbing the erythrocyte/granulocyte pellet. MNCs were washed once with PBS for injection or analysis.

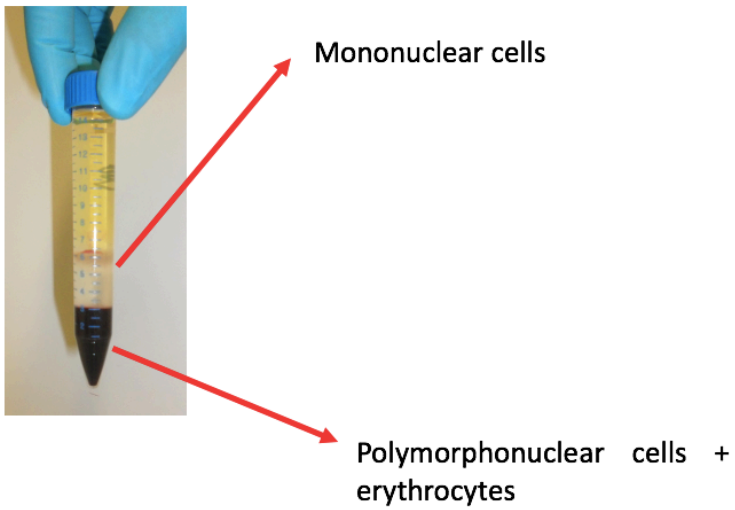


Figure 2.1. Gradient centrifugation of biopsy samples

Liquid biopsy samples from patients underwent gradient centrifugation over lymphoprep to isolate the MNCs.

2.3 *In vivo* experiments

2.3.1 Mice

JAX™ NSG Mice (NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ) were obtained from Charles River, UK and housed in individually ventilated cages (IVCs) under sterile pathogen-free conditions at the University of Cambridge, Central Biomedical Services (CBS) animal facilities, Addenbrooke's Hospital, Cambridge. At the end of procedures mice were culled by the Schedule 1 method of cervical dislocation in accordance with Home Office guidelines under project licence numbers 80/2630 (2014-2017) and P4DBEFF63 (2018) and personal licence IB70151AE.

2.3.2 Injection of primary patient samples into NSG mice

Matrigel® Basement Membrane Matrix (Corning) was stored at -20^{0}C and thawed on ice for a minimum of 4 hours before use. Isolated MNCs were counted following trypan blue exclusion (Thermo Fisher Scientific) and 5X10^6 (or the maximum number of cells available if the biopsy sample was limited) were suspended in 200ul of PBS and mixed with 200ul of Matrigel. Cells were kept on ice and 300ul of suspension injected subcutaneously or interperitoneally into NSG mice using a 30 gauge hypodermic needle (VWR). Mice were monitored for health and tumour size by animal technicians daily. When precise tumour volume was part of a study plan, tumours were measured daily by both a technician and myself. Animals were shaved around the tumour area to assist with precision of measurement and callipers used to measure both the length and width of the tumour.

2.3.3 Harvesting of tumours

Once tumours reached 12mm in any dimension, animals were culled and the tumour removed using sterile scissors and tweezers. Tumours were immediately placed in PBS on ice and returned to the lab for disaggregation. Tumours were disaggregated in a petri dish using a 70μM nylon cell strainer (BD Falcon) and the plunger of a 5ml syringe (BD Plastipak), applying light to medium mechanical force. Single cell suspensions were then

counted following trypan blue exclusion and diluted with PBS for serial passage as described above and/or utilised in further experiments.

2.3.4 Fluorescent in-situ Hybridization

Fluorescent in-situ Hybridization (FISH) was carried out to determine the presence of the MYC translocation. Single cell suspensions were collected by centrifugation at 200 X g for 5 minutes and resuspended in Carbonyl's Fixative (3:1 Methanol: Glacial Acetic Acid). Samples were then sent to Cambridge haematopathology & oncology diagnostic service (HODS), Addenbrooke's Hospital. FISH was performed using a Vysis LSI IGH/MYC/CEP 8 Tri-Color Dual Fusion Probe Kit (Abbott Diagnostics), which detects the fusion between IGH and MYC. Results were reported as the percentage of cells positive for IGH-MYC rearrangement as well as the signal pattern.

2.3.5 Tumour Histology

Following extraction of tumours, either the whole tumour or a section was immediately placed in 10% neutral-buffered Formalin (Sigma-Aldrich). After 24 hours the tumour was removed and placed in 70% Ethanol/PBS. Tumours were submitted to the Histology laboratory at the Department of Pathology, University of Cambridge to be made into Formalin Fixed, Paraffin Embedded (FFPE) blocks. Slides were also cut and Haematoxylin and Eosin staining carried out using standard procedures.

2.4 Flow Cytometry

Cells were counted following trypan blue exclusion, washed with PBS and collected by centrifugation at 250 X g for 5 minutes. Supernatant was discarded and cells resuspended in Flow Cytometry buffer (PBS, 0.05% Bovine Serum Albumin (Thermo Fisher) at a concentration of $1X10^6$ cells/ml. Cell suspension (100µl) was added to each well of a 96-well plate and appropriate antibody was added. Volume of antibodies varied depending on the specific antibody titration and all antibodies used and their associated isotype controls are listed in Table 2.2. Cells were incubated for 40 minutes at 4^0C, protected from light before a further wash in PBS and resuspension in 2ml of flow

cytometry buffer in 5ml polystyrene FACS tubes (BD Falcon™) before immediate analysis on a BD Accuri C6 or a LSR Fortessa Flow Cytometer at the NIHR Cambridge BRC Cell Phenotyping Hub, Addenbrooke's Hospital. At least 10,000 events were recorded and data analysed with FlowJo v9.1 Flow Cytometric Analysis Software.

When separation of cells based on cell surface protein expression was required, Fluorescence Activated Cell Sorting (FACS) was carried out on cells stained as described above using a BD FACSAria™ Fusion at the Cell Phenotyping Hub. Cells were suspended at 30 million per ml and passed through a 50µm cell strainer before sorting.

Antibody	Source	Fluorochrome	Isotype Control
ABCG2	Pharminogen	APC	Mouse IgG2b, κ
CD10	Pharminogen	PE	Mouse IgG_1, κ
CD117	Pharminogen	PE-Cy5	Mouse IgG1, κ
CD117	Pharminogen	PE	Mouse IgG1, κ
CD133	Miltenyi Biotec	PE	Mouse IgG2bκ
CD150	Pharminogen	PE	Mouse IgG1, κ
CD184	Biolegend	PE	Mouse IgG2a, κ
CD19	Pharminogen	PE-Cy5	Mouse IgG1, κ
CD20	Pharminogen	FITC	Mouse IgG2b, κ
CD21	Biolegend	FITC	Mouse IgG1, κ
CD24	Biolgened	PE-Cy7	Mouse IgG2a, κ
CD27	Biolegend	APC	Mouse IgG1, κ
CD34	Biologend	PE	Mouse IgG1, κ
CD34	Pharminogen	PE	Mouse IgG1, κ
CD38	Biolegend	FITC	Mouse IgG2a, κ
CD40	Pharminogen	FITC	Mouse IgG_1, κ
CD44	Biolegend	PE	Mouse IgG1, κ
CD45	Pharminogen	PE	Mouse IgG1, κ
CD45	Pharminogen	APC	Mouse IgG1, κ
CD49D	Pharminogen	FITC	Mouse IgG_1, κ
CD49F	Pharminogen	FITC	Mouse IgG2a, κ
CD59	Biolegend	PE	Mouse IgG2a, κ
CD9	Pharminogen	FITC	Mouse IgG1, κ
CD90	Pharminogen	PE	Mouse IgG_1, κ
CD90	Pharminogen	APC	Mouse IgG1, κ
Oct3/4	Pharminogen	PE	Mouse IgG_1, κ

Table 2.2. Details of antibodies used in flow cytometry

2.5 Side Population analysis

SP analysis was carried out on both cell lines and cells isolated from PDXs. Cell lines were cultured in 2% FBS containing media for 7 days before SP analysis whereas SP analysis of PDX samples was performed on cells immediately after extraction from mice. Live cells were counted following trypan blue exclusion and $1X10^6$ live cells suspended in 1 ml of the indicated FACS Buffer in a 1.5ml Eppendorf tube (Table 2.3).

Cell Type	*Media/Buffer*	*Hepes*	*FBS*
Cell Line	RPMI	10mM	2%
PDX	PBS	10mM	2%

Table 2.3. FACS buffer components

Hoechst 33342 dye (Sigma-Aldrich) was added at the appropriate concentration (Table 2.4) and cells incubated at 370C in a thermomixer set to 200rpm in a dark room for the appropriate time (Table 2.4) Controls were prepared following incubation with 100µM verapamil (Sigma-Aldrich) for 20 mins before the addition of Hoechst dye. In order to determine the correct Hoechst dye concentration and incubation time, all cell lines and PDXs were incubated with concentrations of dye ranging from 2-10ug/ml for 60, 90 or 120 mins until a reliable side population, that was eliminated by verapamil, was identified. Following incubation, cells were collected by centrifugation at 250 X g in a centrifuge at 40C and resuspended in ice cold FACS buffer with 1 µg/ml propidium iodide (Sigma-Aldrich). Sample was transferred to 5ml polystyrene FACS tubes (BD Falcon™) and stored on ice before analysis. Analysis was carried out on BD LSRFortessaTM equipped with an Ultraviolet (UV) laser at the NIHR Cambridge BRC Cell Phenotyping Hub. Hoechst 33342, when excited by UV light, emits fluorescence that can be detected in two distinct channels on the flow cytometer: the “Hoechst Blue” (450/50 nm band-pass filter) and the “Hoechst Red” (610/20 nm long-pass filter) channel, while a dichroic mirror (LP600 nm) was used to split the emission wavelengths. At least 50,000 events were recorded and data analysed with FlowJo v9.1 Flow Cytometric Analysis Software.

When separation of MP and SP populations was required, FACS was carried out on the Hoechst stained cells using a BD FACSAria™ Fusion at the Cell Phenotyping Hub. Cells were suspended at 30 million per ml and passed through a 50µm cell strainer before sorting. The BD FACSAria™ Fusion is equipped with a UV laser and the machine was set up as detailed above.

	Hoechst concentration (µg/ml)	Incubation time (mins)
PDX sample		
1	2	120
2	2	120
3	2	120
4	5	120
5	2	120
Cell Line		
Bjab	2	90
DG-75	2	120
Ramos	5	90
Raji	2	120

Table 2.4. Incubation conditions for SP analysis of PDXs and cell lines

2.6 *In vitro* experiments

2.6.1 Transwell assay

Cell lines - Raji, Ramos, Bjab and DG-75, were incubated with Hoechst 33342 at the optimal concentrations and times as determined above and sorted by FACS using stringent gating based on verapamil-treated control samples. Unsorted whole cell populations were also passed through the cell sorter so all populations experienced the same conditions. Live cell populations were counted following trypan blue exclusion and plated in 24 well plates (10,000 cells per well). The transwell assay was set up using a 6.5 mm Transwell® insert (Corning) with 0.4 µm pore polyester membrane insert. MP

cell were placed in the bottom well (400µl) and SP cells (100µl) placed in the well insert. Cells were counted every second day following trypan blue exclusion.

2.6.2 Measurement of cytokine levels in cell culture media

SP and whole cell populations were grown in media for 5 days before the supernatant was isolated following centrifugation at 200 X g for 5 minutes. Cytokine levels in supernatants was quantified by BioLegend's LEGENDplex™ bead-based immunoassay as indicated by the manufacturer's instructions in a 96 well V-bottom plate. Standards were first produced by serial dilution of the supplied Human Immunoglobulin Isotyping Panel Standard Cocktail. This resulted in 8 standards ranging from 0-400,000pg/mL. Assay buffer (25µl) was added to each well followed by 25µl of sample or standard followed by 25µl of capture beads. The plate was then sealed with a plate sealer, wrapped in tin foil and incubated for 2 hours at room temperature while shaking at 800rpm in order to keep beads in suspension. Cells were then collected by centrifugation of the plate at 250 X g for 5 minutes using a swinging bucket rotor with microplate adaptor. Immediately after centrifugation, the supernatant was disposed of by quickly inverting and flicking the plate in one continuous and forceful motion. The plate was washed by dispensing 200µL of 1X Wash Buffer into each well and incubating for one minute before disposal as described above and before addition of 25µL of Detection Antibody to each well, the plate resealed and covered with aluminium foil for 1 hour incubation at room temperature while shaking at 800 rpm. Without washing, 25µL of SA-PE was then added to each well before re-sealing, wrapping in aluminium foil and incubating at room temperature for 30 minutes with shaking at 800rpm. The plate was subject to centrifugation and supernatant discarded as above before washing and repeating the centrifugation process again. Wash Buffer (150µL) was added to each well and the beads resuspended by pipetting. Samples were read on a BD LSRFortessa^{TM} and data analysed using LEGENDplex^{TM} Data Analysis Software.

2.6.3 Exosome isolation

SP cells from Bjab and DG-75 were grown in media for 3 days. Cells were transferred to 15ml Falcon tubes and collected by centrifugation at 1000 X g for 15 minutes at 4°C in a

bench top centrifuge. Supernatant was collected into new 15ml tubes and the cell pellet discarded. Supernatant was subject to centrifugation at 2000 X g for 15 minutes at 4°C in a bench top centrifuge to remove any remaining cells and was then transferred into 50ml polycarbonate tubes with caps (Beckman Coulter). The volume of supernatant was increased to 20ml with PBS. Tubes, buckets and lids combined were weighed and the values recorded to ensure samples positioned opposite one another in the ultracentrifuge were balanced (max 0.1g weigh difference) with PBS was added dropwise to balance weights. Centrifugation was conducted at 17000g (14500rpm) for 15 minutes at 4°C in an Avanti J-20 centrifuge using a JA20 rotor (17000g = 14500rpm). Supernatant was transferred into 38.5ml polycarbonate tubes with caps (Beckman Coulter) and the remaining pellet (apoptotic bodies) was discarded. Following balancing as described above, centrifugation at 100,000 X g for 80 minutes at 4°C (centrifuge + buckets cooled in advance) in an Optima L-100 XP Centrifuge was performed using a SW-28 rotor (100000g = 27000rpm). Supernatant was discarded and the pellet (exosomes) resuspended in 1ml PBS. Sample was transfer to ½ *2 inch ultra-clear tubes (Beckman Coulter) before balancing volumes to 3ml with PBS and centrifugation at 100000 X g for 40 minutes at 4°C in an Optima L-100 XP Centrifuge with a SW-55Ti rotor (100000g = 32700rpm). As much supernatant as possible was carefully removed using a pipette, with the tube held at head height under light to visualise the exosome pellet and to avoid disturbing it.

2.6.4 Determination of Effective Dose (ED)50 values for individual drugs utilised in experiments

BL cell lines were plated in flat-bottom 96-well plates, (Corning) at a density of $0.5X10^6$ cells/ml and before treatment with a range of 8 concentrations of drug as indicated in the results for 72 hours. Methotrexate, Doxorubicin and Vincristine were purchased from Sigma-Aldrich; Dactolisib, JQ1 and Ibrutinib were obtained from MedChemExpress. Drugs were dissolved in dimethyl sulfoxide (DMSO; Sigma-Aldrich) at a concentration of 2mM then stored at -20^{o}C until use. Rituximab was purchased from Roche in solution and was stored at 4^{o}C at a concentration of 10mg/ml.

Cell viability for each condition was determined using the CellTiter-Blue® (CTB) Cell Viability Assay according to manufacturer's instructions. Briefly, reagent was added directly to cells and incubated for 1 hour at 37°C after which fluorescence was measured on a SpectraMax i3 plate reader (Molecular Devices).

Normalized dose-response curves were generated with GraphPad Prism software[364] using non-linear regression methods that employ the following model (Equation 1), with x being defined as dose value concentration in log scale and y being defined as the normalized response starting from 100% down to 0%.

$$y = \frac{100}{(1 + 10^{(\mathrm{x} - \log \mathrm{ED50})})} \qquad \textbf{Equation 1}$$

Other ED values were calculated using the determined ED50 value by employing the following model (Equation 2), with H being defined as hill slope and F as desired value.

$$\mathrm{EDF} = \frac{\mathrm{F}}{(100 - \mathrm{F})}^{(1/\mathrm{H})} * \mathrm{ED50} \qquad \textbf{Equation 2}$$

2.6.5 Determination of synergy

In order to identify potential synergistic interactions of drugs, BL cell lines were plated at 5X10^6 cells/ml in 96-well plates and treated with single agent or combinations of drugs. The CTB assay was utilised to determine viability and the Bliss combination index (BCI) calculated based on the Bliss Independence model[365]. The BCI was calculated with the following equation: $BCI = \frac{Ea + Eb - Ea * Eb}{Eab}$, where Ea indicates the viability effect of drug A, Eb indicates the viability effect of drug B and Eab indicates the viability effect of being treated with drug A and drug B together. BCI < 1 indicates synergism, BCI = 1 indicates additive effect, BCI > 1 indicates antagonism.

2.7 RT-qPCR quantification of gene expression

All kits were used according to the manufacturer's instructions. To compare gene expression levels, total RNA was isolated with the RNeasy Plus Mini Kit (Qiagen) before RNA (1 ug) was reverse transcribed using the ProtoScript II First Strand cDNA Synthesis Kit (NEB). RT reactions were diluted (1/10) and 10 ng equivalent (to input RNA amount) was used as the template DNA for qPCR using Power SYBR Green PCR Master Mix (ThermoFisher) with standard reaction conditions on the QuantStudio™ 6 Flex Real-Time PCR System (ThermoFisher). Primers used are described in Table 2.5. Data were analysed using the double delta Ct (ΔΔCt) method conducted with normalization to GAPDH (ΔCt) before relative expression level comparison to the control sample (ΔΔCt). All qPCR reactions were performed in technical triplicates.

Gene	Forward primer (5’ - 3’)	Reverse primer (5’ - 3’)
MYC	GGCTCCTGGCAAAAGGTCA	CTGCGTAGTTGTGCTGATGT
BCL2	GGTGGGGTCATGTGTGTGG	CGGTTCAGGTACTCAGTCATCC
BCL6	ACACATCTCGGCTCAATTTGC	AGTGTCCACAACATGCTCCAT
GNA13	CCCAAGGAATGGTGGAAACAA	ACCCAGTTGAAATTCTCGACG
ARID1A	CCAGCAGAACTCTCACGACC	CTGAGCGAAGGACGAAGACG
ID3	GAGAGGCACTCAGCTTAGCC	TCCTTTTGTCGTTGGAGATGAC
TCF3	CCGACTCCTACAGTGGGCTA	CGCTGACGTGTTCTCCTCG
ALDH1A1	CTGCTGGCGACAATGGAGT	CGCAATGTTTTGATGCAGCCT
VEGFA	AGGGCAGAATCATCACGAAGT	AGGGTCTCGATTGGATGGCA
PIM1	GAGAAGGACCGGATTTCCGAC	CAGTCCAGGAGCCTAATGACG
CDK4	ATGGCTACCTCTCGATATGAGC	CATTGGGGACTCTCACACTCT
CDC7	AGTGCCTAACAGTGGCTGG	CACGGTGAACAATACCAAACTGA
TERT	CCGATTGTGAACATGGACTACG	CACGCTGAACAGTGCCTTC
SMARCA4	GACCAGCACTCCCAAGGTTAC	CTGGCCCGGAAGACATCTG
SOX2	GCCGAGTGGAAACTTTTGTCG	GGCAGCGTGTACTTATCCTTCT
GAPDH	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG
PTEN	TTTGAAGACCATAACCCACCAC	ATTACACCAGTTCGTCCCTTTC
COL1A1	GAGGGCCAAGACGAAGACATC	CAGATCACGTCATCGCACAAC
PAX5	ACTTGCTCATCAAGGTGTCAG	TCCTCCAATTACCCCAGGCTT
MTOR	TCCGAGAGATGAGTCAAGAGG	CACCTTCCACTCCTATGAGGC

Table 2.5. RT-qPCR primers used to measure expression of genes in PDXs at different passages

All primers were retrieved from PrimerBank[366,367] *and synthesized by Sigma-Aldrich.*

Gene	Forward primer (5’ - 3’)	Reverse primer (5’ - 3’)
NANOG	TTTGTGGGCCTGAAGAAAACT	AGGGCTGTCCTGAATAAGCAG
ABCG2	ACGAACGGATTAACAGGGTCA	CTCCAGACACACCACGGAT
ABCB1	TTGCTGCTTACATTCAGGTTTCA	AGCCTATCTCCTGTCGCATTA
OCT4	GGGAGATTGATAACTGGTGTGTT	GTGTATATCCCAGGGTGATCCTC
SOX2	TACAGCATGTCCTACTCGCAG	GAGGAAGAGGTAACCACAGGG
AKT1	AGCGACGTGGCTATTGTGAAG	GCCATCATTCTTGAGGAGGAAGT
MCL1	GTGCCTTTGTGGCTAAACACT	AGTCCCGTTTTGTCCTTACGA
PIM1	GAGAAGGACCGGATTTCCGAC	CAGTCCAGGAGCCTAATGACG
FZD6	GATTTTGGTGTCCAAGGCATCT	GCGCACACTGGTCAATTCC
BMI1	CGTGTATTGTTCGTTACCTGGA	TTCAGTAGTGGTCTGGTCTTGT
SOX2	GCCGAGTGGAAACTTTTGTCG	GGCAGCGTGTACTTATCCTTCT
GAPDH	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG

Table 2.6. RT-qPCR primers used to measure expression of genes in SP and MP from PDXs.

All primers were retrieved from PrimerBank[366,367] *and synthesized by Sigma-Aldrich.*

2.8 RNA-sequencing analysis

2.8.1 Library preparation and sequencing

All kits were employed according to the manufacturer’s instructions. Total RNA was extracted from control and relapse tumours using the RNeasy Plus Mini Kit (Qiagen). RNA was quantified by Qbit and submitted to the NGS library facility, Wellcome Trust - Medical Research Council Stem Cell Institute, University of Cambridge for library preparation. RNA integrity for each sample was determined using an Agilent Bioanalyzer. High quality total RNA was verified with a RNA Integrity Number (RIN) value of >9.2 for each sample. mRNA was enriched from total RNA using NEXTflex™ Poly(A) Beads (Bioo Scientific) and libraries then produced using the NEXTflex™ Rapid Directional RNA-Seq Kit (Bioo Scientific), which is designed to prepare directional, strand specific RNA libraries for sequencing using Illumina® sequencers. Briefly, this process involved fragmentation of mRNA which was subjected to first-strand cDNA synthesis with random hexamers. AMPure XP bead clean-up was performed and barcoded DNA

adapters were ligated to both ends of the double-stranded cDNA followed by PCR amplification. Library size was then validated using an Agilent Bioanalyzer and libraries submitted to the CRUK Cambridge Institute Genomics Core where they were multiplexed and single-end sequenced (1 X 50 base pairs) on one lane of an Illumina HiSeq 4000. Samples were demultiplexed based on sample specific barcodes and the raw sequencing data produced.

2.8.2 Data Analysis

Quality Control was performed on the raw reads using FastQC[368] and high quality sequencing data were uploaded to the Galaxy web platform, with the public server at usegalaxy.org used to analyse these data[369]. HISAT2[370], a sensitive spliced alignment program, was used to map RNA-seq reads to the latest version of the human genome, GRCh38, with a minimum alignment rate of 93% for each sample. The BAM files produced were then input into HTSeq-count[371] along with a feature file in GFF format, and a table of counts for each gene was produced as the script output. DESeq2[372] was then utilised to determine differential gene expression (DGE). This method takes count tables generated from HTSeq-count as input and tests for differential expression based on a model using the negative binomial distribution with an output of differentially expressed genes with corresponding Log_2FC, *p* and FDR values produced. Hierarchical clustering and principal component analysis (PCA) were performed using ClustVis[373], a web tool for visualizing clustering of multivariate data with normalised read counts from DESeq2 utilised as the input. In order to carry out Gene Set Enrichment Analysis (GSEA; Broad Institute)[374] normalized read counts were re-formatted to Gene Cluster Text (GCT) file format and a Categorical class (CLS) file was generated (http://software.broadinstitute.org/cancer/software/gsea/wiki/index.php/Data_formats). Both files were used as the input files for GSEA on its javaGSEA Desktop Application. GSEA was performed to compare the expression profile of relapse and control samples using all gene sets from the Molecular Signatures Database (MSigDB) with the msigdb.v6.2.symbols.gmt file downloaded from the GSEA website.

Gene	Forward primer (5’ - 3’)	Reverse primer (5’ - 3’)
MALAT1	TCACCAGCGGAAAACTCAAG	GCTGTAGGCCCCAAAGATGTC
COL1A1	GTGCGATGACGTGATCTGTGA	CGGTGGTTTCTTGGTCGGT
ABCB1	TTGCTGCTTACATTCAGGTTTCA	AGCCTATCTCCTGTCGCATTA
CDC25B	GGCTGAGGAACCTAAAGCCC	CTTTCCGTCTACTGTCTGTAGGA
NFATC4	CTTCTCCGATGCCTCTGACG	CGGGGCTTGGACCATACAG
HDAC10	CAGTTCGACGCCATCTACTTC	CAAGCCCATTTTGCACAGCTC
CDK10	CGGAGGCTCAGGTCAAGTG	CACACAACCCTTGTCGGTCA
MAD2L2	ACCGCCCAGTGGAGAAATTC	CATCGCACACGCTGATCTT
STXBP6	GGATTCGGCAGAGTTTGATTTGT	CCTCCCATAATTTTGGATTGGCA
CD9	TTCCTCTTGGTGATATTCGCCA	AGTTCAACGCATAGTGGATGG
CFL1	TTCAACGACATGAAGGTGCGT	TCCTCCAGGATGATGTTCTTCT
TGFB1	CAATTCCTGGCGATACCTCAG	GCACAACTCCGGTGACATCAA
DBN1	CAACTGGGTGGGCGAAGAT	TGCTGGCGTTCACGATCAC
BID	ATGGACCGTAGCATCCCTCC	GTAGGTGCGTAGGTTCTGGT
PTPRCAP	TTGAGCGACAGGAGGATGAG	GACGCCTCTCCACATTGCT
CD27	CAGAGAGGCACTACTGGGCT	CGGTATGCAAGGATCACACTG
GAPDH	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG

Table 2.7. RT-qPCR primers to validate expression of differentially expressed genes identified from RNA-seq.

All primers were retrieved from PrimerBank[366,367] *and synthesized by Sigma-Aldrich.*

2.9 Genome-wide Cas9 transcriptional activation screen

For genome-wide Cas9 transcriptional activation screens, a three plasmid (Figure 2.2 **B-D**) system - human CRISPR activation library v1 (Addgene, cat. no. 1000000057, kind gift from Feng Zhang, MIT) - was used[375].

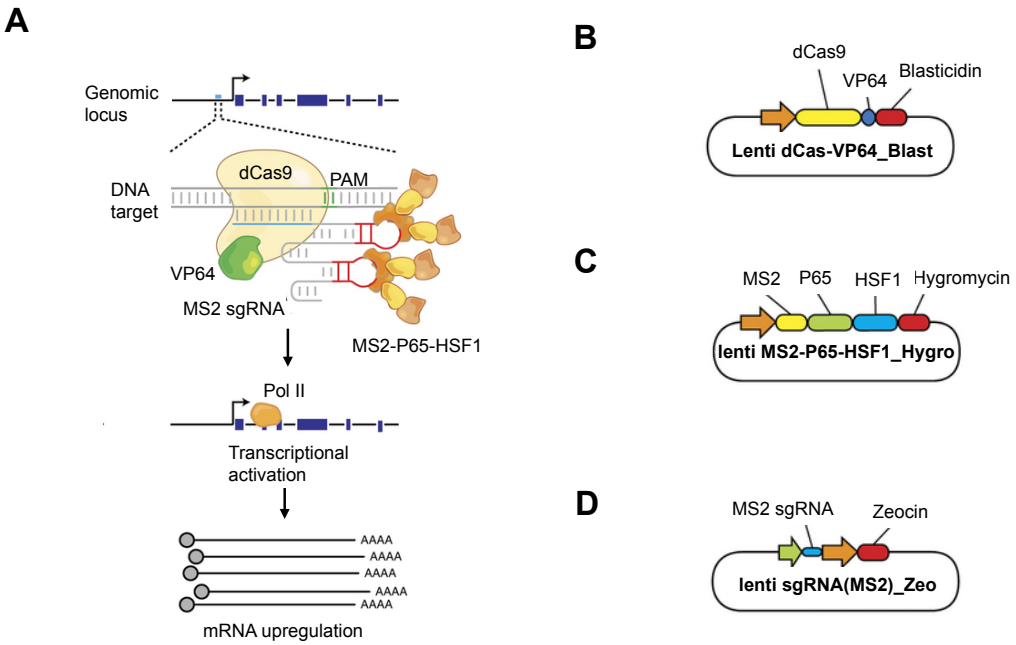


Figure 2.2. Genome-scale Cas9 transcriptional activation screen employed to identify resistance mechanisms to chemotherapy and Rituximab

*(**A**) Programmable transcriptional activation can be achieved by using dCas9 and activation domains (e.g., VP64/p65/HSF1) to recruit transcriptional machinery to the transcriptional start site (TSS) of the desired target gene, resulting in upregulation of the target transcript. PAM: protospacer adjacent motif, Pol II: RNA polymerase II. (**B-D**) Vector-system used for the Genome-scale Cas9 transcriptional activation screen. The library (lenti sgRNA(MS2)_Zeo) must be combined with additional SAM effectors in a 3-vector format. Blast: Blasticidin, Hygro: Hygromycin, Zeo: Zeocin. Figure modified from Joung et al., 2017*[376].

This library is based on a dCas9-VP64 fusion protein (Figure 2.B) that recruits transcriptional complexes to the transcriptional start site (TSS) of target transcripts as well as an altered sgRNA (Figure 2D) to recruit accessory transcriptional co-activators (MS2-p65-HSF1) (Figure 2C) to synergistically interact with the transcriptional complex. Specifically, a hairpin aptamer that selectively binds to MS2 phage protein was appended to the tetraloop and stem loop no. 2 regions[377]. Then a separate vector (Figure 2C) was constructed to express MS2 protein fused to the p65 transcription factor and heat shock transcription factor 1 (HSF1). This allows for additional recruitment of transcriptional activators to the TSS, leading to >500-fold enhanced overexpression in targeted mRNA levels compared to dCas9-VP64 alone[375].

2.9.1 Generation of a dCas9 and MS2 expressing Raji cell line

To conduct the genome-scale Cas9 transcriptional activation screen, a stably expressing lenti dCAS-VP64_Blast (Addgene, cat. no. 61425) and lenti MS2-P65-HSF1_Hygro (Addgene, cat. no. 61426) Raji cell line was engineered[375]. Briefly, the lentiviral plasmids (kind gifts from Feng Zhang, MIT) were individually packaged into lentivirus and subsequently used for viral transduction of the Raji cell line. First, the expression plasmids were co-transfected into log-phase growth 293FT cells (Invitrogen) with second-generation lentiviral packaging plasmids (kind gifts from Didier Trono, EPFL) – psPAX2 (Addgene, cat. no. 12260) and pMD2.G (Addgene, cat. no. 12259) at a ratio of 1:1:1 with TransIT-293 (MirusBio) in OptiMEM reduced serum medium (Thermo Fisher Scientific) at 32°C. After 24 hours, the medium was replaced with fresh Dulbecco's Modified Eagle Medium (DMEM; Gibco) supplemented with 10% FBS, DNAse I (Thermo Fisher Scientific, 1 U/mL), 20 mM HEPES pH 7.4 and 5 mM $MgCl_2$. Viral supernatant was collected 48 and 72 hr post-transfection, pooled, subject to centrifugation at 360 x g for 5 min, the resulting supernatant filtered (Minisart 0.45µm filter, Sartorius Stedim) and stored at -80^{o}C if not used immediately.

Following these procedures, Raji cells were transduced with viral particles of lenti dCAS-VP64_Blast at a Multiplicity of Infection (MOI) of < 0.7 then selected in blasticidin S HCl (5 µg/mL) (Thermo Fisher Scientific) 24 hrs post viral transduction for seven days or until 0% survival was observed in uninfected cells. Then lenti MS2-P65-HSF1_Hygro particles at an MOI of < 0.7 were applied and stable cells were selected in hygromycin B (300 µg/mL) (Thermo Fisher Scientific) 24 hrs post viral transduction for seven days as before. Cells were maintained in blasticidin/hygromycin B-containing media thereafter at the above concentrations.

2.9.1.1 Western Blot to confirm the presence of dCas9

Protein lysates were prepared by incubating 5 million cells in RIPA buffer (Sigma Aldrich, cat. no. R0278) supplemented with Halt Protease and Phosphatase Inhibitor Cocktails (ThermoFisher) for 30 min on ice. Cellular debris were removed by centrifugation at 16,000 X g for 20 min at 4°C and the supernatant stored at -20°C until usage. Proteins were quantified with a Pierce™ BCA Protein Assay Kit (ThermoFisher) following the manufacturer's protocol and analyzed on a SpectraMax i3 plate reader (Molecular Devices). Protein lysate (50 µg) was solubilized in SDS solution (250 mM Tris pH 6.8, 10% SDS, 5% beta-mercaptoethanol (BME), 0.02% bromophenol blue, 30% Glycerol) and boiled for 5 mins at 95°C. Samples were resolved by SDS-PAGE with a 4% stacking gel and a 10% running gel for 1 hr at 100 V in running buffer (25 mM Tris, 190 mM glycine, 0.1% SDS). Proteins were transferred to a 0.45 µM Immobilon-P PVDF membrane (Merk Millipore) using a Mini Trans-Blot Cell apparatus (Bio-Rad) at 100 V for 1 hr in transfer buffer (25 mM Tris, 0.01% SDS, 1.92 mM Glycine). Following transfer, membranes were first blocked in buffer (5% milk powder in TBST (20 mM Tris-HCl pH 8, 150 mM NaCl, 0.1% Tween-20)) for 1 hr at room temperature and subsequently incubated with polycloncal anti-Cas9 antibody (Abcam, 1:2000 dilution in blocking buffer) at 4°C overnight. Afterwards the membrane was incubated with goat anti-rabbit immunoglobulin/HRP (Agilent technologies, 1:2000) for 1 hour at room temperature. Washing of membranes was performed in TBST (3 X 5 minute washes), protein bands were visualized by Immobilon Western Chemiluminescent HRP Substrate (Merck Millipore) and detected by a LAS-4000 Image Analyzer (Fujifilm/Raytek).

2.9.1.2 Validation of dCas9-VP64 overexpression activity

To confirm the dCas9-VP64 overexpression activity in the transduced Raji cell line, 30 guide sequences (Table 2.8) were designed with the Cas9 activator tool[378] and introduced into the lenti sgRNA(MS2)_Zeo backbone vector (Addgene, gift from Feng Zhang, MIT). The genes – encoding for growth factors, guanine nucleotide exchange factors, transcription factors, kinases and adapter proteins were targeted with two sgRNAs per gene.

Gene	sgRNA	Forward ssDNA element (5' - 3')	Reverse ssDNA element (5' - 3')
ARHGEF9	1	CACCgCGCAAGCGCCCGCAGTCGCT	AAACAGCGACTGCGGGCGCTTGCGc
	2	CACCgGGGCGGGTATGTCAGTGGCTC	AAACGAGCCA`CTGACATACCCGCCc
BCL10	1	CACCgAGGTACTGACAAGCCCAGAC	AAACGTCTGGGCTTGTCAGTACCTc
	2	CACCgCAGGGTCTGGGAAAAGGGG	AAACCCCCTTTTCCCAGACCCTGCc
CRK	1	CACCgACGCGCTCCCTGCCCAGGG	AAACCCCTCGGGCAGGGAGCGCGTc
	2	CACCgGGCTGCCGGGAAGGGGCCCG	AAACCGGGCCCCTTCCCGGCAGCCc
EREG	1	CACCgGGTGCTGCGAACTTTATACT	AAACAGTATAAAGTTCGCAGCACCc
	2	CACCgGAGCCCCTCCCGGGCCCTAA	AAACTTAGGGCCCGGGAGGGGCTCc
FOS	1	CACCgAAACGTCACGGGCTCAACCA	AAACTGGTTGAGCCCGTGACGTTTc
	2	CACCgGGCGCCAGAGGGGTGGCGCG	AAACCGCGCCACCCCTCTGGCGCCc
NR4A2	1	CACCgCAGCGCGGCGATTGGGCGGC	AAACGCCGCCCAATCGCCGCGCTGc
	2	CACCgGGGCCAGGAGTCCAGGGAGCG	AAACCGCTCCCTGGACTCCTGGCCc
PLEKHG6	1	CACCgCAGTAATGAGGGCTGAGTAA	AAACTTACTCAGCCCTCATTACTGc
	2	CACCgGGGCGGGCGCCGCCGGGGG	AAACCCCCCGGCGGCGCCCCGCCCc
PRKACA	1	CACCgTGCGGGGGCGTCACAGACAG	AAACCTGTCTGTGACGCCCCCGCAc
	2	CACCgGGCCTAGGCCAATGAGCGGC	AAACGCCGCTCATTGGCCTAGGCCc
PSD2	1	CACCgGGGACGGACGGCAGGAGGGA	AAACTCCCTCCTGCCGTCCGTCCCc
	2	CACCgGGAGCGGAGCCGTGAGCTGG	AAACCCAGCTCACGGCTCCGCTCCc
PTGES	1	CACCgCTGCAGGGAAAGCACAAAGT	AAACACTTTGTGCTTTCCCTGCAGc
	2	CACCgGGGCGGTGCTGGCTGCAGGA	AAACTCCTGCAGCCAGCACCGCCCc
RORC	1	CACCgCCCCCAGTGCTTCTGGACTG	AAACCAGTCCAGAAGCACTGGGGGc
	2	CACCgGTTTAAGCTGTGCTGCACCACAC	AAACGTGTGGTGCAGAGCTTAAACc
SAMD4A	1	CACCgCTCCCCCCTTCTTGCAGCTT	AAACAAGCTGCAAGAAGGGGGGAGc
	2	CACCgGATGGTGATTTCCGGCGTCC	AAACGGACGCCGGAAATCACCATCc
SPDEF	1	CACCgTATAATGGGAAATCAGGCCC	AAACGGGCCTTGATTTCCCATTATAc
	2	CACCgTTTGTTCAGGTAAATAAGGA	AAACTCCTTATTTACCTGAACAAAc
UBIAD1	1	CACCgGCCGGCGGCGCGGGCTGGAC	AAACGTCCAGCCCGCCGCCGGCc
	2	CACCgCCGCCCTCCAGCCCACCCTC	AAACGAGGGTGGGCTGGAGGGCGGc
YAP1	1	CACCgGGCGAGTTTCTGTCTCAGTC	AAACGACTGAGACAGAAACTCGCCc
	2	CACCgCAAACGCCAAAACTAAAGTT	AAACAACTTTAGTTTTGGCGTTTGc
NT		CACCgGGTCCCTCAGGGTGCAACTT	AAACAAGTTGCACCCTGAGGGACCc

Table 2.8. Oligonucleotides used to generate dsDNA fragments containing the 20 bp target sequence.

Previously cloned into the lenti sgRNA(MS2)_zeo vector. ARHGEF9: RhoA guanine nucleotide exchange factor 9 (TF), BCL10: B-Cell CLL/Lymphoma 10 (GF), CRK: CRK Proto-Oncogene (Adapter Protein), EREG: Epiregulin (GF), FOS: Fos Proto-Oncogene (TF), NR4A2: Nuclear Receptor Subfamily 4 Group A Member 2 (TF), PLEKHG6: Pleckstrin Homology And RhoGEF Domain Containing G6 (GNEF), PRKACA: Protein Kinase CAMP-Activated Catalytic Subunit Alpha, PSD2: Pleckstrin And Sec7 Domain Containing 2, PTGES: Prostaglandin E Synthase (GF), RORC: RAR Related Orphan Receptor C (nucleic acid binding protein), SAMD4A: Sterile Alpha Motif Domain Containing 4A (TF), SPDEF: SAM Pointed Domain Containing ETS Transcription Factor, UBIAD1: UbiA Prenyltransferase Domain Containing 1, YAP1: Yes Associated Protein 1 (TF). NT is a scrambled 20 bp sequence that is predicted to not target any region of the human genome. Orange nucleotides indicate the overhangs while black nucleotides represent the 20 bp target sequence.

Following this, the cloned lenti sgRNA(MS2)_zeo backbone was transduced into library-ready stably expressing lenti dCAS-VP64_Blast and lenti MS2-P65-HSF1_Hygro Raji cells and selected for integration of the vector in zeocin-containing media (300 µg/mL) (Invivogen). Total RNA was collected seven days after the final transduction with the RNeasy Plus Mini Kit (Qiagen) and quantified by Nanodrop 1000 (ThermoFisher) with RT-qPCR carried out as described in section 2.7. All qPCR reactions were performed in technical triplicates. RT-qPCR primers used are listed in Table 2.9.

Gene	Forward primer (5' - 3')	Reverse primer (5' - 3')
ARHGEF9	AATGAGCACTGAGCGTCACTA	AGCAGGGTCCTATCTCGCTG
BCL10	TCTGGACACCCTTGTTGAATCT	TGGAAAAGGTTCACAACTGCTAC
CRK	GCGGAGTAGCTGGTACTGG	GCGCGAGTTCTCTGAGACG
EREG	GTGATTCCATCATGTATCCCAGG	GCCATTCATGTCAGAGCTACACT
FOS	GGGGCAAGGTGGAACAGTTAT	CGGCTTGGAGTGTATCAGTCA
NR4A2	ACCACTCTTCGGGAGAATACA	GGCATTTGGTACAAGCAAGGT
PLEKHG6	CCGCCCTACAGAAGCTGAAG	GGATAATGGTCGAGAACTCAGGA
PRKACA	ACCCTGAATGAAAAGCGCATC	CGTAGGTGTGAGAACATCTCCC
PSD2	GGATGGCCTGTCAGACTCAGA	CAGCCTGCTAAACTCGTTGTT
PTGES	TCCTAACCCTTTTGTCGCCTG	CGCTTCCCAGAGGATCTGC
RORC	GTGGGGACAAGTCGTCTGG	AGTGCTGGCATCGGTTTCG
SAMD4A	TCGAGGCTTTGGGCAATCC	GAGCTGACGAATCCACTGGT
SPDEF	CAGTGCCCGGTCATTGACA	CAGCCGGTATTGGTGCTCT
UBIAD1	AGTGTGCCTCCTACGTGTTG	CAGGACACCGTGGGATCTG
YAP1	TAGCCCTGCGTAGCCAGTTA	TCATGCTTAGTCCACTGTCTGT
GAPDH	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG

Table 2.9. RT-qPCR primers used for the validation of the dCas9-VP64-system based overexpression of target genes in the Raji cell line

All primers were retrieved from PrimerBank[366,367] *and synthesized by Sigma-Aldrich.*

2.9.2 CRISPR screen of Chemotherapy and Rituximab resistance

2.9.2.1 Transformation, amplification and preparation of lentiviral sgRNA libraries

Aliquots of the Human CRISPR Activation Library v1 (~1000 ng at 50 ng/µl) (SAM - 3 plasmid system) (Addgene) were acquired. Each library (100 ng) was transformed into ElectroMAX Stbl4 Competent Cells (ThermoFisher) via electroporation with the Gene Pulser II Electroporation System (BioRad) according to the protocol supplied with the Stbl4 cells. This process was carried out with 4 replicate reactions. Cells were recovered in SOC Outgrowth Medium (NEB) and cultured for 1.5 hrs at 30 °C before being plated onto ampicillin lysogeny broth (LB) agarose bioassay dishes (Nunc), which were incubated overnight at 30 °C. The bacterial colonies were then harvested and the CRISPR plasmid libraries were isolated using the EndoFree Plasmid Maxi Kit (Qiagen).

To prepare lentivirus libraries, log-phase HEK293FT cells (Invitrogen) were seeded at ~50% confluence the day before transfection in reduced (5%) FBS supplemented DMEM to obtain ~80% confluence at transfection. For each T175 flask, 18 µg of sgRNA plasmid library, 16 µg of psPAX2 and 10 µg of pMD2.G were incubated in 4.5 mL of Opti-MEM/TransIT-293 (110 µL of TransIT-293) for 15 min at room temperature. After transfection (60 hrs), the media were collected and subject to centrifugation at 500 X g, 4°C for 15 mins to remove the cellular debris. Aliquots of the supernatant were stored at –80 °C.

2.9.2.2 Transduction of Raji cells using lentiviral CRISPR libraries

To determine optimal virus volumes for achieving a MOI of 0.3, trial infections to determine effective MOI were set up with Raji cells. Briefly, 1 million cells per well were seeded into 6-well plates (Corning) in 10% FBS RPMI 1640 and the virus libraries mixed with fresh media were added to the cells. Each well received a different volume of the viruses (10 µl up to 500 µl) along with a mock (no virus) transduction control. Cells were incubated with the virus for 24 hours and cells from each well were then split into 2 wells in a 6-well plate. One well received 300 µg/mL of zeocin. After 5 days, when no viable cells remained in the no-transduction control, live cells in duplicate wells

containing zeocin were counted following trypan blue exclusion. The effective MOI was calculated as the average cell count from the triplicate wells (with zeocin selection) divided by the average cell count from the triplicates with no selection reagent. The virus volume yielding a MOI of 0.3 was selected for scaling up for the screens. The virus transduction protocol for the actual screens was scaled up to 20 million cells per T175 flask with a total of 500 million cells being transduced. Cells were selected in zeocin and the effective MOI of the cells transduced for the screens were verified once again as described above.

2.9.2.3 SAM (GOF) enrichment screen

Raji cells (2 billion) were transduced as described above to maintain ~500 X representation of the sgRNA libraries. Zeocin was added to the cells 24 hr after transduction and maintained for 7 days replacing the selection media every 2 days. On day 7, media on the cells was replaced with media not containing zeocin and maintained for 3 days before dividing them into four different conditions in duplicates with a minimum of 50 million cells per replicate and a separate 50 million cells frozen down in duplicate to be used as the day 0 sample for genomic DNA analysis. The four conditions employed were: Chemotherapy treated cells, Rituximab treated cells, Chemotherapy + Rituximab treated cells and the DMSO control. ED concentrations used were: Chemotherapy = ED70, Rituximab = ED70, Chemotherapy + Rituximab = ED85 which were determined based on the results of the synergy experiments detailed in chapter 7. Cells were either counted and divided or had their media replaced every 2-3 days, depending on their confluence. Cell pellets with a minimum of 50 million cells per replicate were taken 13 days after the addition of drugs at which point the screen was terminated and the cell pellets were frozen at -80 °C.

2.9.2.4 Preparation of HiSeq libraries

Frozen cell pellets were lysed and their gDNA extracted with the QIAamp DNA Blood Maxi Kit (Qiagen) according to the manufacturer's instructions. PCR of the virally integrated guides was performed on gDNA at the equivalent of 500 cells per guide using Herculase II Fusion DNA Polymerase (Agilent) (Table 2.10). Oligonucleotides (Sigma)

included the Illumina adapters, a staggered region, and the barcodes (Table 2.11). The amount of input gDNA for each sample to achieve 500 X coverage of the SAM library was calculated to be approximately 230 µg DNA per replicate, assuming 6.6 µg of gDNA is the yield of 1 million cells. Therefore, 23 parallel 100 µl PCR reactions with 10 µg of input genomic DNA in each reaction were completed in an ABI Veriti Thermal Cycler (Applied Biosystems) in a single-step reaction of 24 cycles (Table 2.12).

Component	Amount per reaction (µl)	Final concentration
UltraPure water	61	
5 x Buffer	20	1x
100% DMSO	2	2%
100 mM dNTP	1	1 mM
10 µM F Primer Mix	2.5	0.25 µM
10 µM R Primer	2.5	0.25 µM
Herculase II Fusion DNA Polymerase	1	
1 µg/mL gDNA	10	0.1 µg/mL

Table 2.10. PCR conditions for amplification of virally integrated guides.

Primer	Sequence (5’ - 3’)
F1	**AATGATACGGCGACCACCGA**GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**TAAGTAGAGG***CTTTATATATC**TTGTGGAAAGGACGAAACACC*
F2	**AATGATACGGCGACCACCGA**GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**ATCATGCTTA***GCTTTATATAT**CTTGTGGAAAGGACGAAACACC*
F3	**AATGATACGGCGACCACCGA**GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**GATGCACATCT***GCTTTATATA**TCTTGTGGAAAGGACGAAACACC*
F4	**AATGATACGGCGACCACCGA**GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**CGATTGCTCGAC***GCTTTATAT**ATCTTGTGGAAAGGACGAAACACC*
F5	**AATGATACGGCGACCACCGA**GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**TCGATAGCAATTC***GCTTTATA**TATCTTGTGGAAAGGACGAAACACC*
F6	**AATGATACGGCGACCACCGA**GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**ATCGATAGTTGCTT***GCTTTAT**ATATCTTGTGGAAAGGACGAAACACC*
F7	**AATGATACGGCGACCACCGA**GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**GATCGATCCAGTTAG***GCTTTA**TATATCTTGTGGAAAGGACGAAACACC*
F8	**AATGATACGGCGACCACCGA**GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**CGATCGATTTGAGCCT***GCTTT**ATATATCTTGTGGAAAGGACGAAACACC*
F9	**AATGATACGGCGACCACCGA**GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**ACGATCGATACACGATC***GCT**TTATATATCTTGTGGAAAGGACGAAACACC*
F10	**AATGATACGGCGACCACCGA**GATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**TACGATCGATGGTCCAGA***GC**TTTATATATCTTGTGGAAAGGACGAAACACC*
R1	**CAAGCAGAAGACGGCATACGAGAT**GAAGAAGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT***GCCAAGTTGA**TAACGGACTAGCCTT*
R2	**CAAGCAGAAGACGGCATACGAGAT**CGTTACCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT***GCCAAGTTGAT**AACGGACTAGCCTT*
R3	**CAAGCAGAAGACGGCATACGAGAT**GTCTGATG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT***GCCAAGTTGA**TAACGGACTAGCCTT*
R4	**CAAGCAGAAGACGGCATACGAGAT**TTACGCAC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT***GCCAAGTTGAT**AACGGACTAGCCTT*
R5	**CAAGCAGAAGACGGCATACGAGAT**TTGAATAG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT***GCCAAGTTGA**TAACGGACTAGCCTT*
R6	**CAAGCAGAAGACGGCATACGAGAT**TCCTTGGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT***GCCAAGTTGAT**AACGGACTAGCCTT*
R7	**CAAGCAGAAGACGGCATACGAGAT**ACAGGTAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT***GCCAAGTTGA**TAACGGACTAGCCTT*
R8	**CAAGCAGAAGACGGCATACGAGAT**AGGTAAGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT***GCCAAGTTGA**TAACGGACTAGCCTT*
R9	**CAAGCAGAAGACGGCATACGAGAT**AACAATGG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT***GCCAAGTTGA**TAACGGACTAGCCTT*

Table 2.11. Oligonucleotides used for HiSeq library preparation.

All forward oligonucleotides were mixed at equal molar concentrations, introducing artificial diversity to the library from the varied staggered sequences, then used with one reverse primer (unique 8 bp molecular barcode) per sample. ***Illumina P5 adaptor****,* ***Illumina P7 adaptor****, sequencing forward,* ***sequencing reverse****,* ***stagger (varied) sequence****, index barcode, priming site.*

Cycle number	Denature	Anneal	Extend
1	**95 °C, 20 sec**		
2-25	**95 °C, 20 sec**	**60 °C, 20 sec**	**72 °C, 40 sec**
26			**72 °C, 3 min**

Table 2.12. PCR cycle conditions used to amplify the sgRNAs' guide sequence region and to append the Illumina (HiSeq) compatible adapters and barcodes.

PCR products from all 23 reactions were pooled, purified using the Zymo DNA Clean and Concentrator-5 kit (Zymo research) and a small amount of sample separated on a 2% agarose gel in 1 x TAE at 100 V for 2.5 hrs to confirm the removal of excess primers. The isolated amplicons were then quantified by qPCR reactions using kapa Kit (ThermoFisher), mixed at equal molar concentrations and submitted for HiSeq high Output v4 x 100 reactions with PhiX spike at Bauer Core (Harvard University).

2.9.2.5 Bioinformatics-based analysis of CRISPR screen results

Raw fastq files were downloaded from the Bauer Core's server using FileZilla. Initial QC was conducted with FastQC (Babraham Bioinformatics) to assess the general quality of the sequence runs. The following steps were executed on Terminal of a MacBook Air laptop. First, multiple fastq files for each of the replicates were merged into a single fastq file with the concat function. All nucleotides preceding the 20bp guide sequence of the sgRNAs were removed using CutAdapt 1.12[379]. Resulting merged and trimmed files were used as the input files for analysis on a turn-key bioinformatics script, named the Model-based Analysis of Genome-wide CRISPR/Cas9 Knockout (MAGeCK)[380]. A subsequent version of the script, MaGECK-visualization framework (VISPR) was used to generate the figures directly associated with the screen results[381].

2.9.2.6 Viability assay based validation of candidates identified from the screen

To validate the top candidates identified from the screen, 2 guide sequences per candidate gene (Table 2.13) were individually cloned into the aforementioned sgRNA vector. Once cloned and the sequence verified, each were co-transfected with psPax2 and pMD2.G into 1 million 293FT cells per individual well of a 6-well plate using TransIT-293. Viruses were collected 60 hours post-transfection and frozen at -80°C.

CRISPR-ready Raji cells (with stable expression of dCas9-VP64 + MS2-cofactors) were transduced with the sgRNA containing viruses to generate individual stable cell lines for each sgRNA. Cells were selected (300 μg/ml zeocin) and passaged for 5 days before splitting them into 96-well plates for drug treatment and CTB-based viability assay. Each stable cell line was plated at $5X10^6$ cells/ml and treated with concentrations of chemotherapy and Rituximab as utilised in the screen. Further analysis of the shift in ED50 induced by each sgRNA for single drugs was performed by treating cells with 5 concentrations of drug for 3 days and determining the ED50 as outlined in section 2.6.4.

Gene	sgRNA	Sequence (5' - 3')	Sequence (5' - 3')
ABCC1	1	CACCGCAGCCGGACCAGCCACCTCT	AAACAGAGGTGGCTGGTCCGGCTGC
	2	CACCGCGCCCCACCCCGCCCCGCCC	AAACGGGCGGGGCGGGGTGGGGCGC
PIM1	1	CACCGCGGGACTGGGCGACTCCCCT	AAACAGGGGAGTCGCCCAGTCCCGC
	2	CACCGGGCCGAGGGAGGAGAGGGGG	AAACCCCCCTCTCCTCCCTCGGCCC
IRAK1	1	CACCGGAGGGGGCGCGAGGGACAGC	AAACGCTGTCCCTCGCGCCCCCTCC
	2	CACCGGCGCAAGATGGCGGCGGGCC	AAACGGCCCGCCGCCATCTTGCGCC
PEG10	1	CACCGCTGAACTGGCCAAGCTGGGA	AAACTCCCAGCTTGGCCAGTTCAGC
	2	CACCGGGGGATGGGGGATGGGGGTG	AAACCACCCCCATCCCCCATCCCCC
BRCC3	1	CACCGGGGAAACGCCCGCGCGGTTT	AAACAAACCGCGCGGGCGTTTCCCC
	2	CACCGGGGCTGTGGGGGCGGGGCTG	AAACCAGCCCCGCCCCCACAGCCCC
ATG10	1	CACCGGGTGGTCGCGGGCGGCGGGG	AAACCCCCGCCGCCCGCGACCACCC
	2	CACCGGCGGGGGAGAGCTGCCCGCC	AAACGGCGGGCAGCTCTCCCCCGCC
NFKB1	1	CACCGCTAGGGGCGCTGTGCGCCCC	AAACGGGGCGCACAGCGCCCCTAGC
	2	CACCGGTGCGGGCTTCCCCCACCCC	AAACGGGGTGGGGGAAGCCCGCACC
KRAS	1	CACCGCGTACGAGAGGGAGCGGCTG	AAACCAGCCGCTCCCTCTCGTACGC
	2	CACCGGCGGTGTGGGAAGAGGGAAG	AAACCTTCCCTCTTCCCACACCGCC
ALCAM	1	CACCGGTTAGTCCCAGGGCTCTCCC	AAACGGGAGAGCCCTGGGACTAACC
	2	CACCGGTTGCAGCGCCCCAGTGATA	AAACTATCACTGGGGCGCTGCAACC
ITGB1	1	CACCGAGCACTTAAAGCCGCCGGCC	AAACGGCCGGCGGCTTTAAGTGCTC
	2	CACCGCCGGGGCTAGGAGGAGGCGG	AAACCCGCCTCCTCCTAGCCCCGGC
ABCB1	1	CACCGCACATATTTCTTCAATGCTT	AAACAAGCATTGAAGAAATATGTGC
	2	CACCGGGATAAGTTTGGGTGGAGGA	AAACTCCTCCACCCAAACTTATCCC
LIN7B	1	CACCGCAGTGGGAGGGGTCGCCAAC	AAACGTTGGCGACCCCTCCCACTGC
	2	CACCGGGGCGCGGCTGGCGCGCGGC	AAACGCCGCGCGCCAGCCGCGCCCC
CITED2	1	CACCGCCGGAACACCGCGAACACCC	AAACGGGTGTTCGCGGTGTTCCGGC
	2	CACCGCGTCTTGGCGCCGCCCGCCC	AAACGGGCGGGCGGCGCCAAGACGC
MAPK13	1	CACCGGAAGCGCGGGAGCGCAGGGC	AAACGCCCTGCGCTCCCGCGCTTCC
	2	CACCGAGGCCCGGGCGTGGCGGCCC	AAACGGGCCGCCACGCCCGGGCCTC
SOX14	1	CACCGTCTGCTGCTGATTGGCGCCC	AAACGGGCGCCAATCAGCAGCAGAC
	2	CACCGTGTTGGCAGGTGCGCCCCCT	AAACAGGGGGCGCACCTGCCAACAC
ADAM9	1	CACCGCACCGCTAGTGCTGGTTGCCACCGC	AAACGCGGTGGCAACCAGCACTAGC
	2	CACCGCACCGGTCGGGCGCGCGTGCTCGTC	AAACGACGAGCACGCGCGCCCGACC
DDB2	1	CACCGCACCGAAAAAAAATCCATAAAGCCG	AAACCGGCTTTATGGATTTTTTTTC
	2	CACCGCACCGTGACCTCCCTGGAGCAAAGA	AAACTCTTTGCTCCAGGGAGGTCAC
DKK3	1	CACCGAGAGGAGGGCGTGGGGCGTG	AAACCACGCCCCACGCCCTCCTCTC
	2	CACCGGCGGGGCGGGGTGGGCCTGG	AAACCCAGGCCCACCCCGCCCCGCC
JARID2	1	CACCGCGGTCGGCACCAGGCTAAGC	AAACGCTTAGCCTGGTGCCGACCGC
	2	CACCGGCGCCTGGCGGCTGAGCGGT	AAACACCGCTCAGCCGCCAGGCGCC
HAX1	1	CACCGTTTAGGAGGAAGGCGGTCAT	AAACATGACCGCCTTCCTCCTAAAC
	2	CACCGAAACTGGTCTCTGAAAGGTG	AAACCACCTTTCAGAGACCAGTTTC

Table 2.13. Oligonucleotides used to introduce guide sequences into sgRNA vectors for the genes identified to induce resistance in the SAM screen against chemotherapeutics and/or Rituximab.

2.10 Statistical Analysis

This project involved several different approaches of statistical analysis and all statistical tests were performed with Prism 7, unless otherwise indicated.

Chapter 3. Establishment of Patient Derived Xenografts

3.1 Introduction

The standard toolkit for cancer researchers for decades has been the use of clonal cell lines cultivated *in vitro.* However, this resource is lacking in its ability to allow for investigation of heterogeneity within a patient's tumour. Data has highlighted that the behaviour of these cell lines diverges largely from the original tumour. In particular, gene expression patterns show significantly reduced complexity in cultured cell lines as forced selection by *in vitro* conditions reduces the heterogeneity of the sample with subsequent loss of important biological properties[382]. Furthermore, mouse xenografts using these cell lines have had minimal success when it comes to predictive power in translational research, which contributes to the high failure rate of most novel therapies in phase III clinical trials[383]. Patient derived Xenografts (PDXs) have evolved as a powerful pre-clinical tool capable of bridging the gap between established cell lines and primary tumour samples. PDXs better maintain the heterogeneity of patient tumours and hence allow for a more clinically relevant examination of tumour evolution, response to therapy and development of chemo-resistance.

Mature paediatric B-NHL (pB-NHL) presents in three main forms - BL, DLBCL and PMBL[384] and despite advances in the understanding of the different pathologic and genetic features of these separate entities the current standard of care for all is still high dose chemotherapy with the addition of Rituximab being currently formally evaluated, with positive outcomes[102]. Within the most common subtype, BL, recent studies have shown that it may not be as homogeneous as previously thought, i.e. there are distinct subpopulation with unique genotypes and phenotypes within tumours. Love *et al.* found, in the largest sequencing study of BL, that there was considerable heterogeneity in the number of mutated genes across samples. Out of a total of 70 mutated genes the range of mutation in the sample cohort was 2 to 16, highlighting diversity across patient samples[52]. This study did not report variant allele frequency of mutations although research in our own lab, carrying out high coverage multi-region whole exome sequencing has identified a large amount of low frequency, sub-clonal heterogeneity within primary presentation BL samples. Given this evidence of heterogeneity,

preclinical model systems such as PDXs that capture the genetic and functional heterogeneity of primary pB-NHL are necessary. This will enable us to enter the next stage of research where we try to identify targeted therapies for individual patients as well as decipher the clonal dynamics leading to resistant disease in 5-10% of patients. PDX models of B cell lymphoma exist, although those published are adult samples and consist mainly of the DLBCL subtype as this is most common in adults[385]. To our knowledge no other comprehensive pB-NHL PDX resource exists.

Of course, the PDX model does not come without its challenges; studies of PDXs have indicated that histology[386] and gene expression profiles are retained[387], along with single-nucleotide polymorphism[388] and copy number variants[389], although conflicting data exits regarding the presence of genetic alterations in the engrafted tumours compared with their parental cancers[390–392]. Genes associated with stromal gene ontology annotations are the most altered, most likely due to loss of the human stromal compartment and interactions with the mouse stromal microenvironment. For these reasons, careful investigations at each passage are important to ensure the PDX is faithfully recapitulating the original tumour sample. Most authors advocate using PDX models with a low passage number (<10) to preserve the genetic integrity of the parental tumor[393], although most PDX models in use have not been extensively genetically profiled, something that should become common practice going forward. Another consideration when establishing PDX models is the location of implantation with subcutaneous (SC), intraperitoneal (IP), intravenous (IV) and orthotopic implantation all commonly used. Given the "semi-solid" nature of pB-NHL careful consideration is necessary to determine the most relevant and reliable site for the establishment of PDXs in this study. Numerous types of immunodeficient mice can be used to establish xenograft models - athymic nude mice, SCID, NOD-SCID, and recombination-activating gene 2 (Rag2)-knockout mice,[303] although primary malignant cells often require more robust immunodeficiency for effective engraftment. The use of NOD/SCID mice with IL2rg mutations (NSG) has proven to be effective across a range of cancers including those with lymphoid origin and was chosen for use in this model. Of course, the use of NSG mice results in replacement of human stromal components (such as cancer-associated fibroblasts, endothelial cells, immune and inflammatory cells) by

murine elements as well as a lack of interaction between immune cells and tumour cells. Transgenic mouse models would provide these tumour-stromal-immune interactions but the fact that they do not represent a human origin is a major limiting factor. For example, transgenic MYC mouse models show delayed onset of lymphomagenesis and result in tumours with of a clonal nature eliminating the ability to examine tumour heterogeneity[59]. The utilisation of humanised PDX models in the future, if costs reduce, will help to combine all the important elements of tumour modelling but for now the use of PDXs in NSG mice offers the best method to examine heterogenous populations of pB-NHL.

Chapter Aims

- Establish Patient Derived Xenografts of samples obtained from the Inter B-NHL Ritux Clinical Trial for paediatric B-cell non-Hodgkin lymphoma
- Determine the optimal engraftment route
- Determine the ability of the PDX model to maintain tumour characteristics through passage with respect to histology, genetic features, cell surface protein expression and gene expression.

3.2 Results

3.2.1 Patient biopsy sample details

Due to the rarity of pB-NHL, ethics constraints and standard practice to take fine needle aspirates for diagnosis, acquiring samples of the required quality and quantity for PDXs has been difficult (i.e. fresh frozen disaggregated tumour or effusion). Samples were received from various biopsy sites including a solid tumour mass, pleural effusion, bone marrow and peripheral blood (Table 3.1).

Patient No.	Biopsy site	Diagnosis	Stage
1	Needle biopsy of abdominal mass	Burkitt Lymphoma	III
1	Bone Marrow	Burkitt Lymphoma	III
2	Pleural Effusion	Burkitt Lymphoma	III
3	Peripheral blood	Burkitt Lymphoma/Leukaemia with CNS involvement	IV
3	Bone Marrow	Burkitt Lymphoma/Leukaemia with CNS involvement	IV
4	Peripheral Blood	Burkitt Lymphoma	IV
4	Bone Marrow	Burkitt Lymphoma	IV
5	Bone Marrow	Burkitt Lymphoma/Leukaemia	IV
6	Bone Marrow	Burkitt Lymphoma	III
7	Pleural Effusion	Burkitt Lymphoma	III
8	Bone Marrow	Originally Burkitt Leukaemia. Changed to Pre-B-ALL with MYC translocation	IV
9	Bone Marrow	Burkitt Lymphoma/Leukaemia	IV
9	Pleural Effusion	Burkitt Lymphoma/Leukaemia	IV
10	Peripheral Blood	Burkitt Lymphoma	IV
10	Bone Marrow	Burkitt Lymphoma	IV

Table 3.1. Site of biopsy and diagnosis for each patient sample received.

3.2.2 Successfully engrafted tumours maintain the surface phenotype of biopsy tumour cells

PDXs were successfully established from pleural effusion, bone marrow and peripheral blood in this study. All samples were treated identically upon receipt with mononuclear cells (MNCs) isolated by gradient centrifugation over lymphoprep. When sufficient biopsy material was available cells were analysed by flow cytometry using forward scatter (FSC) and side scatter (SSC) to identify the MNC population reminiscent of lymphocytes. In addition, CD20, CD19 and CD10 expression was assessed and the main population in which these cells resided was used as an indicator of the tumour cell population. B cell surface proteins are indicative of the presence of tumour cells, specifically CD20 which is a membrane-embedded surface molecule that plays a role in the development and differentiation of B cells into plasma cells. It appears after CD19 and CD10 expression which are other characteristic markers in BL[394]. Some interesting observations regarding engraftment and maintenance of tumour cell characteristics were noted in successfully engrafted tumours. The first PDX was produced from a pleural effusion of patient 2. Flow cytometry to determine cell size and granularity identified a small population of cells (approximately 5%) consistent with a lymphoid cell profile (Figure 3.1A). Following isolation of the MNCs by gradient centrifugation over lymphoprep, this increased to 55% (Figure 3.1B). These cells had a B-NHL phenotype, with over 99% being CD10, CD19 and CD20 positive, a population not usually present in pleural effusion (Table 3.2). Interestingly, surface proteins associated with stemness were also noted within this CD20 population including CD34 and ABCG2.

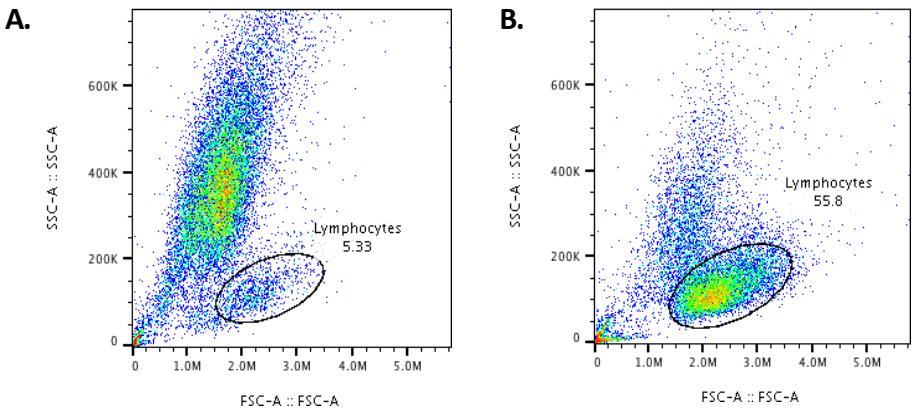


Figure 3.1. Gradient centrifugation of a pleural effusion from patient 2 increases the ratio of lymphoid cells detected by flow cytometry.

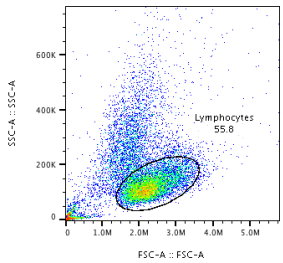
(A). Whole pleural effusion from the biopsy was analysed by flow cytometry. The lymphocyte gate was approximately 5% of the whole population, based on the normal FSC and SSC for this population. (B). MNCs were separated from the pleural effusion by gradient centrifugation and this results in an increased lymphocyte gate of 55.8%

Marker	%
CD5	**0**
CD9	**0**
CD10	**99.7**
CD19	**99.8**
CD20	**99.3**
CD21	**24.9**
CD24	**99.7**
CD27	**32.6**
CD34	**1.36**
CD38	**99.8**
CD40	**15**
CD44	**89.7**
CD45	**99.7**
CD49D	**44**
CD49F	**0**
CD59	**83.6**
CD90	**0.7**
CD105	**0**
CD133	**0**
CD117	**0**
CD135	**0**
CD150	**0**
CD184	**99.7**
ABCG2	**1.13**

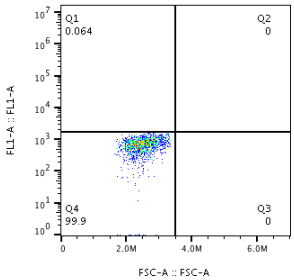
Table 3.2. Cell surface phenotype of lymphocytes within the pleural effusion taken from patient 2.

Given the identification of a clear population with a B-NHL phenotype the possibility of sorting this population before injection into mice was considered. However, due to the limited sample size as well as the potential that FACS sorting may induce stress in these primary patient samples, the whole MNC population was injected into mice. Furthermore, the presence of accompanying healthy cells may provide a microenvironment more similar to that of the patient and hence could be beneficial for initial growth of the tumour cells in mice. This method of sample preparation led to successful tumour engraftment after SC injection of mice with cells ($5X10^6$) in matrigel. Importantly, phenotyping of the tumour produced at passage 1 (P1) highlighted the maintenance of the cell surface profile of tumour cells present in the biopsy (Figure 3.2).

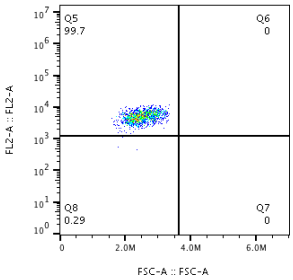
Pre-engraftment



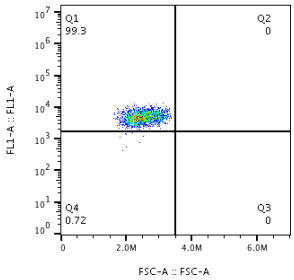
CD9



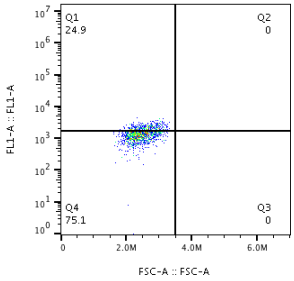
CD10



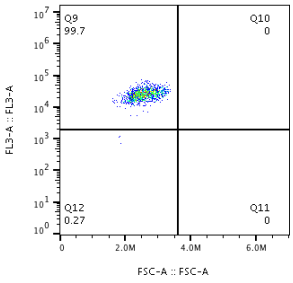
CD20



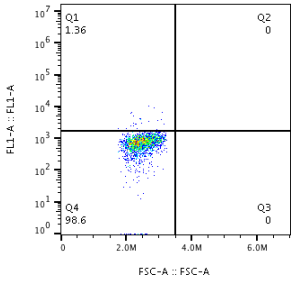
CD21



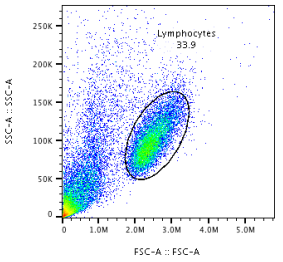
CD24



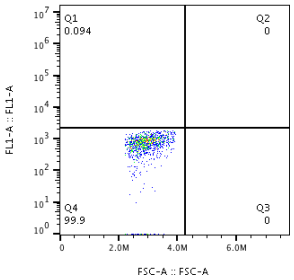
CD34



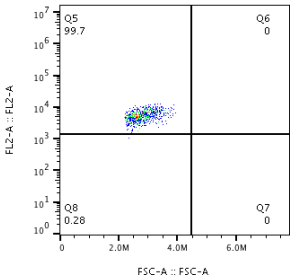
After Passage



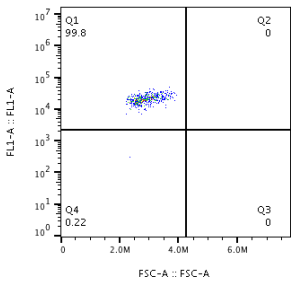
CD9



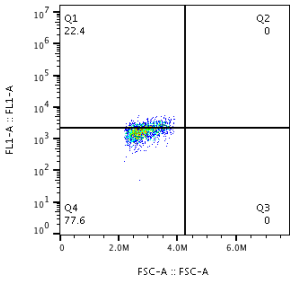
CD10



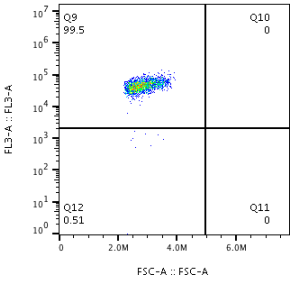
CD20



CD21



CD24



CD34

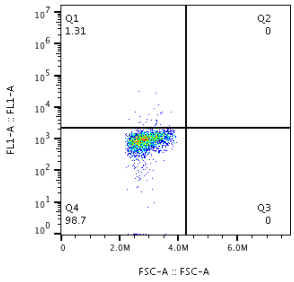


Figure 3.2. The cell surface phenotype of all markers is maintained in patient sample 2 between biopsy and passage 1.

MNCs were isolated from pleural effusion sample and cells analysed by flow cytometry. The lymphocytes were gated on as shown in the FSC and SSC graph. Cells were single stained for high, low and medium expressed surface markers. Following injection of tumour cells in mice and extraction of the tumour produced this process was repeated and the expression profiles compared, showing maintenance of marker expression. Data is representative of 2 replicates as biopsy cells were injected into 2 different mice.

The second successful engraftment was of a peripheral blood sample from patient 3. Following gradient centrifugation to isolate MNCs, cells were labelled with anti-CD20 and a back-gating strategy identified the population with highest CD20 expression. (Figure 3.3B). The percentage of CD20 positive cells constituted 84% of the pre-determined lymphocyte gate of the primary sample (Figure 3.3C) which increased to 96% following successful engraftment. The percentage of cells expressing other BL associated surface proteins also increased following the first passage (Table 3.3) suggesting that *in vivo,* malignant $CD20^+$ cells are selected for growth over other cells within the sample such as untransformed lymphocytes. In contrast, the percentage of $CD117^+$ cells disappears after the first passage suggesting that either these cells were not tumour cells or were not selected for growth *in vivo.* In support of the former, these cells were negative for CD20 in the biopsy sample (Figure 3.4A). Furthermore ABCG2 expression was maintained on 1% of the population after passage and this cell population was positive for CD20 in the initial biopsy (Figure 3.4B).

Results of the first passage of patient samples 2 and 3 highlight some noteworthy observations. The presence of healthy cells within the injected population does not hamper engraftment and these untransformed cells do not survive. The surface protein expression profile of tumour cells is maintained after passage in the mouse, including populations present at very small percentages. Finally, cell surface proteins associated with stemness such as CD34 and ABCG2 were present in the biopsy samples and were maintained in the PDX model.

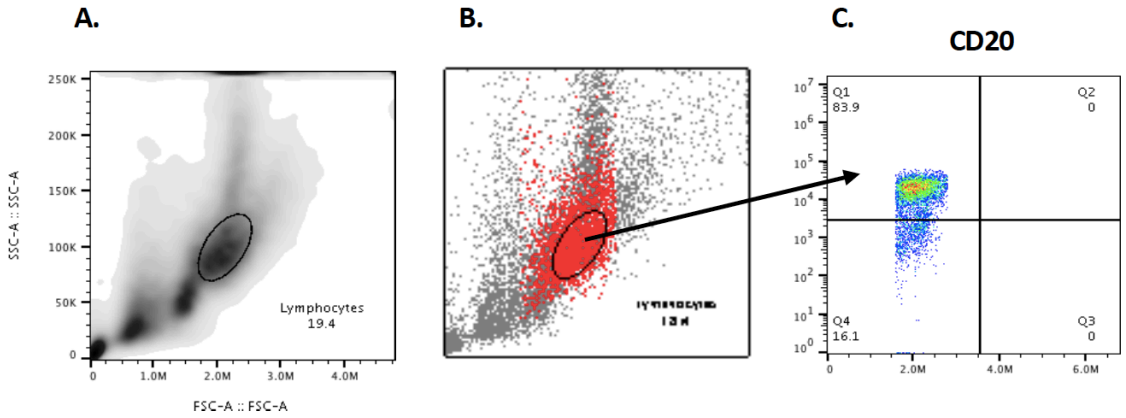


Figure 3.3. Gating strategy to identify lymphocyte gate with CD20 positive cells

Peripheral blood sample underwent gradient centrifugation to isolate MNCs. (A). Cells were analysed by flow cytometry and the FSC and SSC plot identified three clear populations of cells. (B.) The lymphocyte gate was chosen based on the highest expression of CD20 (red cells). (C). 83.9% of cells in this gate were CD20 positive.

Marker	Biopsy	Passage 1
CD9	63.7	89.4
CD10	60.6	94.3
CD19	79.5	99.1
CD20	83.9	96.1
CD34	5.97	0.82
CD38	90	98.6
CD40	3.11	0.5
CD44	3.3	0.6
CD45	83.7	96.4
CD49D	19.7	92.4
CD49F	0.96	0
CD90	0.08	0
CD105	0.06	0
CD133	1.37	0
CD117	1.1	0
CD21	0.68	0
CD24	80.3	89
CD27	82	98
CD59	93	99.5
CD135	0	0
CD150	0.17	0
ABCG2	1.98	1
CD184	91.8	99.2

Table 3.3. Cell surface phenotype of lymphocytes from patient sample 3 peripheral blood before engraftment and after P1 highlights the preferential survival of certain cells

MNCs were isolated from peripheral blood sample and single stained for an array of surface markers before analysis by flow cytometry. Following injection of tumour cells in mice and extraction of the PDX this process was repeated and the expression profiles compared. Data is representative of 2 replicates as biopsy cells were injected into 2 different mice.

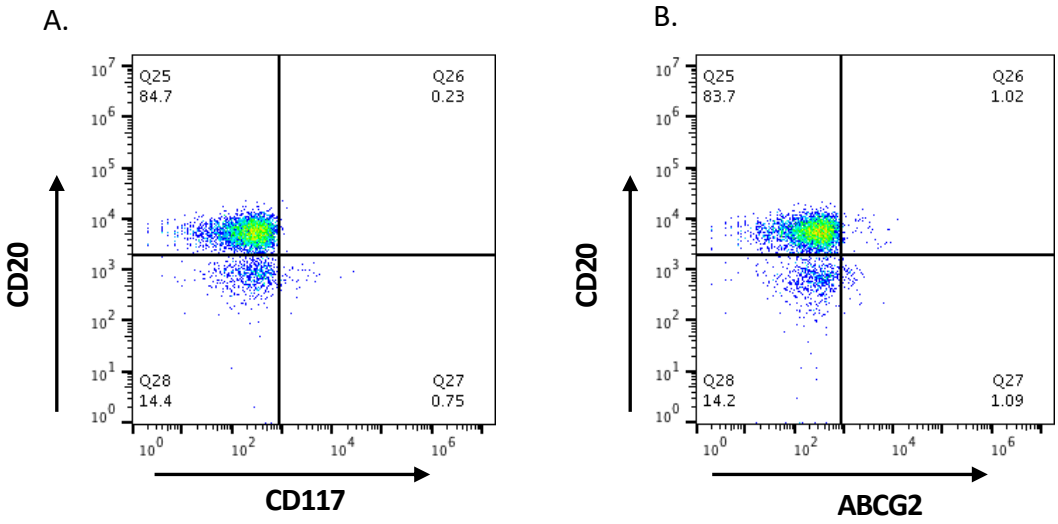


Figure 3.4. $CD20^{+}CD117^{+}$ cells were not present in patient 3 biopsy sample but $CD20^{+}ABCG2^{+}$ cells were

MNCs were isolated from peripheral blood and double stained for (A). CD20 + CD117 and (B). CD20 + ABCG2. Cells were then analysed by flow cytometry and the presence of double positive cells in the biopsy sample determined.

3.2.3 Not all biopsy samples led to successful engraftment

Successful PDXs were formed from pleural effusion, bone marrow and peripheral blood during this study, suggesting that the site of biopsy may not be important in the establishment of PDXs from BL samples. Despite this, not all biopsies gave rise to tumour formation. In fact, the only biopsy sample received from a solid mass (patient 1) did not engraft. Bone marrow with tumour involvement was also received from this patient and also did not lead to successful engraftment. Importantly, the patient had not received any treatment before the biopsy was taken. The tumour biopsy was limited in size hence all cells were injected, although analysis was carried out on the bone marrow sample. Three clear populations of MNCs were present after collection by centrifugation; gates L1, L2 and L3 (Figure 3.5, Table 3.4). L1 had no CD20 positive cells while L2 had the highest proportion of CD10, CD19 and CD20 positive cells and hence it was deemed most likely the tumour cells lay within this gate. The failure of engraftment for both biopsy

samples from this patient suggests that the environment in which cells were injected was not appropriate for growth of these particular tumour cells. An overview of engraftment outcome for each biopsy sample received is given in Table 3.5. In most cases sufficient material was available to perform a limited amount of phenotyping on the original biopsy and the percentage of cells expressing BL associated markers (CD10, CD19 and CD20) was noted and gave an indication of tumour cell burden in the sample.

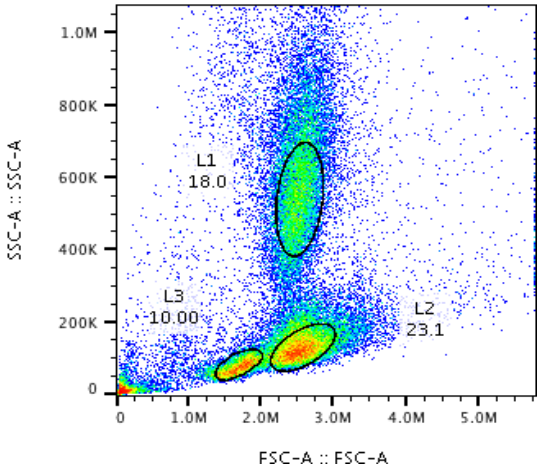


Figure 3.5. Multiple cell populations were present within the isolated MNC fraction from bone marrow of patient sample 1

Following gradient centrifugation of patient sample 1 bone marrow biopsy, MNCs were isolated and analysed by flow cytometry. Forward and side scatter identified three distinct populations within the MNCs.

Marker	L2 %	L3 %
CD5	**0.82**	**61.3**
CD7	**0.28**	**59.1**
CD9	**13.2**	**9.25**
CD10	**77.6**	**3.74**
CD19	**90.3**	**20.1**
CD20	**35.2**	**19.1**
CD34	**0.62**	**0.26**
CD38	**93.7**	**28.5**
CD40	**36.3**	**4.65**
CD44	**10.7**	**82**
CD45	**59.7**	**93.9**
CD49D	**6.76**	**20.2**
CD49F	**0.71**	**11.7**
CD90	**0**	**0.16**
CD105	**0.31**	**0**
CD133	**0.31**	**0.38**
OCT3/4	**0**	**0**
CD117	**0.46**	**0.2**

Table 3.4. Cell surface expression of the indicated proteins on distinct cell populations within the bone marrow of patient 1.

Patient	1	1	2	3	4	4	5	6	7	8	9	9	10	10
Biopsy material	Tumour mass	Bone Marrow	Pleural Effusion	Peripheral Blood	Peripheral Blood	Bone Marrow	Bone Marrow	Bone Marrow	Pleural Effusion	Bone Marrow	Pleural Effusion	Bone Marrow	Peripheral Blood	Bone Marrow
CD10		77.6	99.3	60.7	0.3	5.1	5.7	54.9	99.7		10.5	8.33	82	64.3
CD19		90.3	96.7	79.5	2.8	9.9	16.6	69.9	99.9		7.23	13.89	67.8	59.2
CD20		35.2	95.1	83.7	5.2	6.6	17.9	42.1	96.2		8.38	7.07	79	56.9

Table 3.5. Engraftment rate of biopsy samples

Biopsy site and presence of cells with a BL phenotype is reported for each sample received. Green represents biopsy samples that lead to successful engraftment and growth of a PDX. Red represents engraftment failure. All CD20 positive cells are also positive for expression of CD10 and CD19.

It is difficult to form any conclusive insights from the above data regarding factors that are important for engraftment of BL PDXs. It appears pleural effusion is the biopsy material most likely to engraft with three out of three injected samples forming PDXs. Bone marrow and peripheral blood each produced one successful PDX but other samples from these biopsy sites did not engraft. The tumour cell burden may also be an important factor in determining if a sample of MNCs engrafts or not. The first four successful engraftments all had a high burden of tumour cells (based on $CD20^{+}CD19^{+}CD10^{+}$ phenotype) although the fifth successful engraftment from a pleural effusion had a relatively low burden of $CD20^{+}CD10+CD19^{+}$ cells and the bone marrow sample from this patient did not produce a PDX, suggesting again that pleural effusion is the material most likely to engraft. Other factors also likely influence engraftment rate. For example, although each biopsy was processed identically, some samples were received more quickly than others and the sub-optimal conditions that may have been experienced by samples may have impacted the probability of engraftment. Nonetheless, five PDXs were formed from a total of ten patients giving a 50% engraftment rate. Alternatively, when considering each of the 14 individual biopsies received across these patients, all of which had tumour involvement, the engraftment rate was 35.7%. A frozen aliquot of a further established BL PDX sample were received from another investigator. This pleural effusion-derived PDX successfully engrafted bringing the total number available to us to six. The PDX from patient 9 was established later in the project and is not included in further studies.

3.2.4 Subcutaneous or intraperitoneal injection of tumour cells but not the intravenous route leads to successful engraftment

When sufficient patient material was available, samples were injected SC, IP and IV. Interestingly, even for samples obtained from bone marrow and peripheral blood, the IV route did not lead to tumour formation or the detectable presence of human cells in the animal peripheral circulation. However, all samples that engrafted SC also developed tumours via the IP route (Figure 3.6) and both these routes led to tumours with the same cell surface phenotype (Figure 3.7).

A.

B.

Figure 3.6. Tumours form from both subcutaneous and intraperitoneal injections

MNCs were isolated from patient sample 7 pleural effusion and suspended in PBS. Cells were injected with matrigel ($5X10^6$) in both the SC and IP route. (A). SC tumours are easy to palpate and measure and form tumours outside the peritoneal membrane. (B). IP tumours form inside the peritoneal cavity without intrusion into organs.

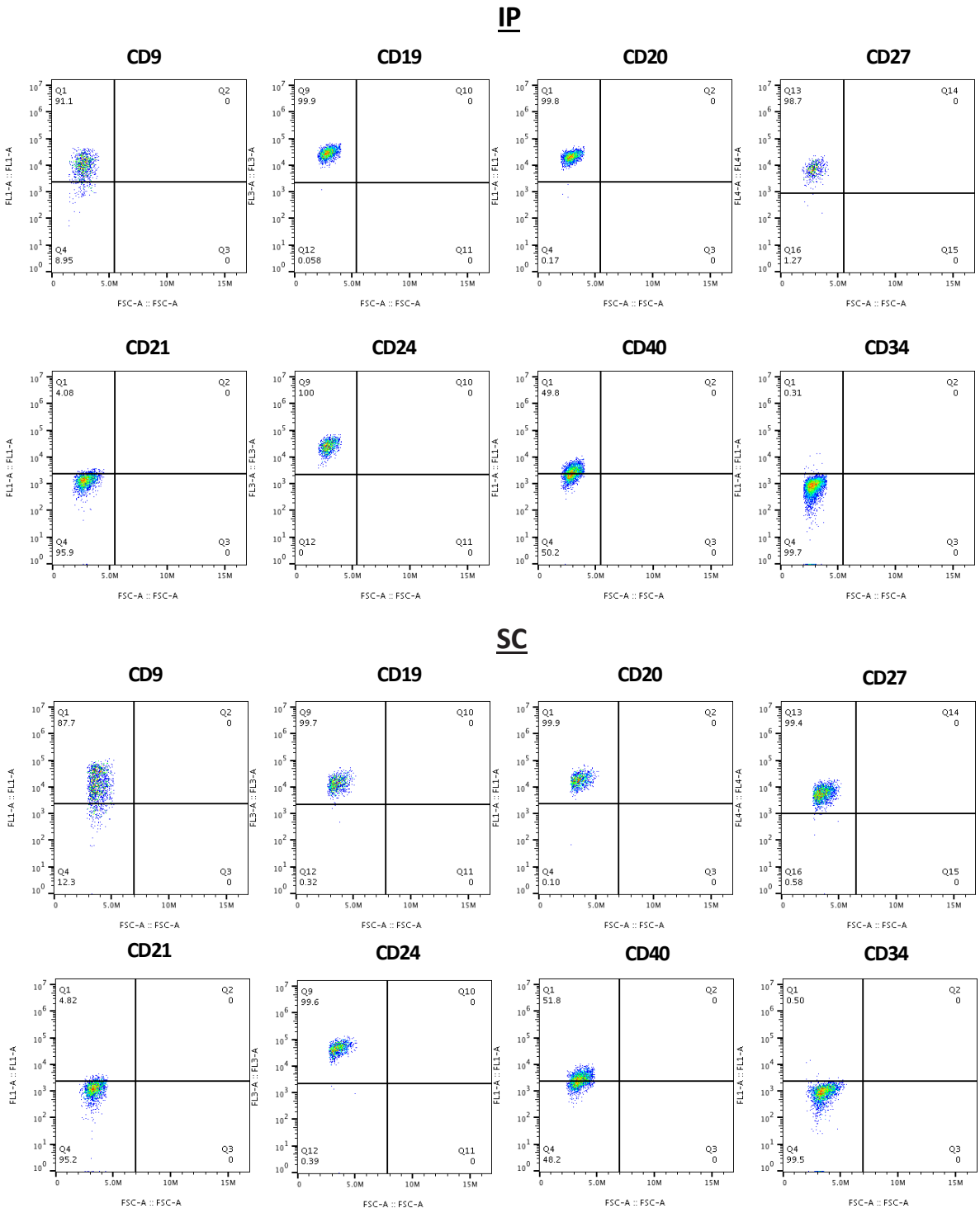


Figure 3.7. Cell surface expression of tumours is the same whether formed in the subcutaneous or intraperitoneal cavity.

MNCs were isolated from patient sample 7 pleural effusion and suspended in PBS. 5X10^6 cells were injected with matrigel in both the SC and IP route. When tumours reached maximum size they were excised, disaggregated into single cell suspension and analysed by flow cytometry for surface protein expression. The same results were seen for the other PDX samples, with no significant difference in phenotype based on SC or IP injection.

The fact that tumour engraftment was successful via the SC and IP route but not IV, highlights the potentially important role matrigel has in establishing tumours in the mouse, possibly due to the range of extracellular matrix proteins and growth factors present. Furthermore, the failure of IV injections may hint towards the importance of tumour cell-to-cell interaction, with matrigel providing a 3D matrix. As tumours propagated by IP and SC routes had no significant phenotypic surface expression variability, the SC method was employed throughout these studies to allow easier measurement and monitoring of tumour growth.

3.2.5 Monitoring of tumour characteristics through passage

To ensure that PDX tumours are a robust and reliable model for studying tumour heterogeneity it is imperative to monitor tumour characteristics through passage to ensure they are maintained. Comprehensive phenotyping by histology, FISH, flow cytometry and gene expression was conducted on all PDXs at passage one and again at subsequent passages.

3.2.5.1 Tumours maintain their histopathological appearance over multiple passages

Due to defective macrophages in NSG mice the starry sky appearance associated with solid BL masses is not fully recapitulated. Despite this, histopathological analysis of H and E stained slides from each tumour after passage 1 and 5 were examined and the maintenance of a high-grade lymphoma histology was verified along with tumour cell size and architecture (Figure 3.8).

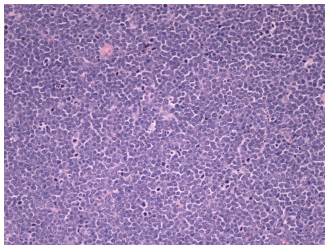
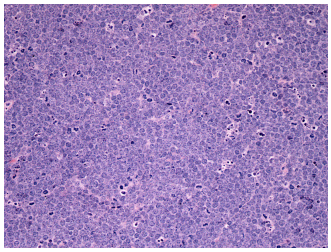
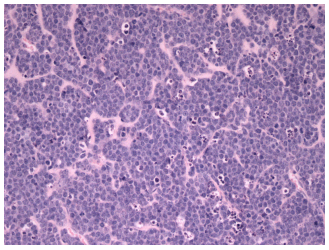
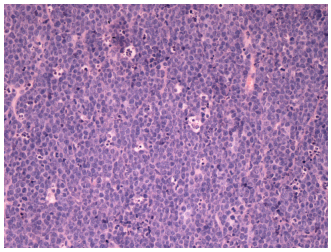
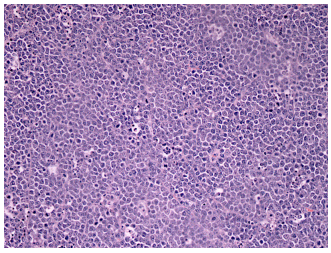
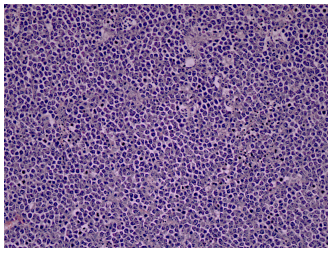
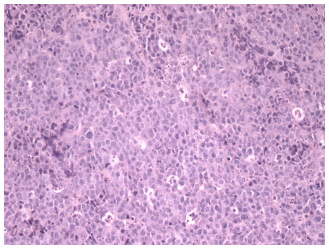
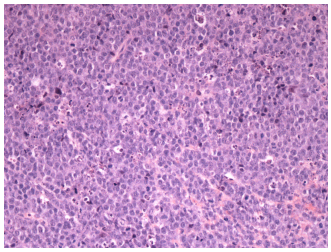
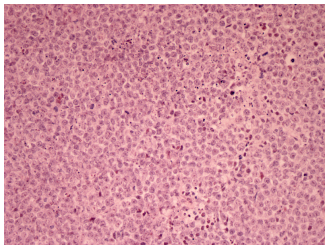
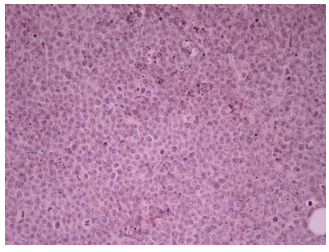
	Passage 1	Passage 5
PDX1		
PDX2		
PDX3		
PDX4		
PDX5		

Figure 3.8. Tumour histology is maintained in PDXs between passage 1 and 5

PDX tumours were extracted when they reached 12mm and formalin fixed. H and E staining was performed and slides examined by a clinical histopathologist.

3.2.5.2 Tumours maintain their cell surface protein expression profile over multiple passages

The average numbers of days for the injected primary cells to first produce palpable tumours when engrafted SC was 33 (range 28-48). Following tumour growth the cells were reinjected as outlined in Figure 3.9. Subsequent passages, for the most part, led to palpable masses of 12mm in any one direction by an average of four weeks. At each passage tumours were analysed for cell surface protein expression and each tumour faithfully maintained its expression profile over passage, highlighting the existence of an intricate population equilibrium within individual samples (Figure 3.10).

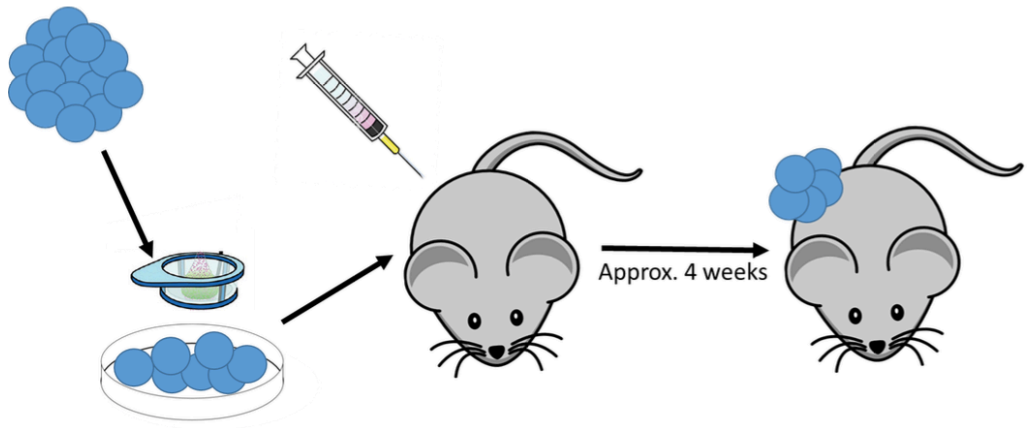


Figure 3.9. Mechanical disaggregation of tumour for serial passage

Once initial engrafted biopsy samples had formed tumours of 12mm, animals were euthanised and tumours extracted. Tumours were mechanically disaggregated into single cell suspension for serial passage in NSG mice.

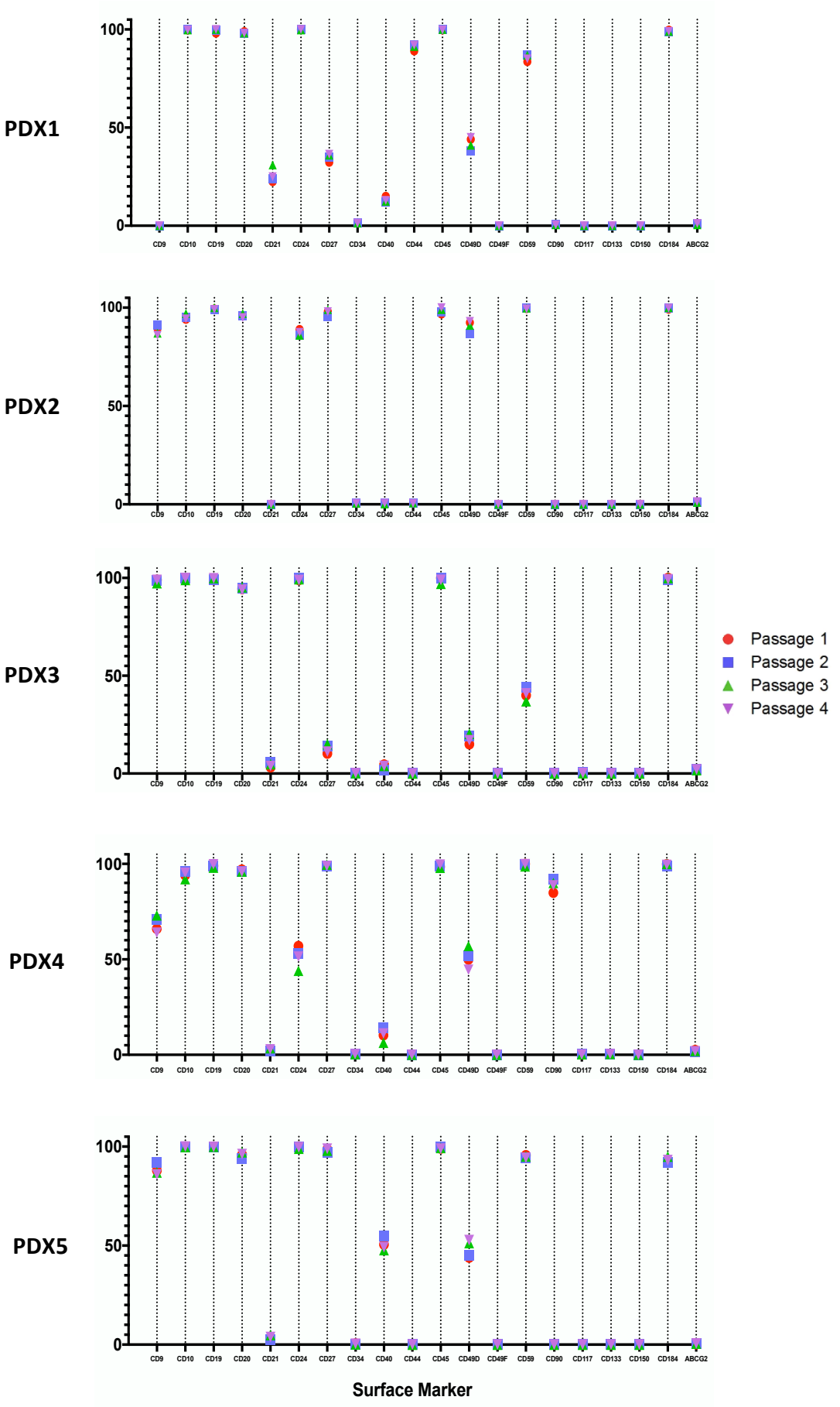

PDX1
PDX2
PDX3
PDX4
PDX5
100
50
0
CD9 CD10 CD19 CD20 CD21 CD24 CD27 CD34 CD40 CD44 CD45 CD49D CD49F CD59 CD90 CD117 CD133 CD150 CD184 ABCG2
Passage 1
Passage 2
Passage 3
Passage 4
Surface Marker

Figure 3.10. Cell surface protein expression of 5 PDXs was maintained over 4 passages.

Biopsy tumour cells were injected into NSG mice with matrigel. When the tumour reached 12mm (Passage 1) it was extracted, disaggregated into single suspension and labelled with antibodies before flow cytometry analysis. Cells were also serially passaged. The process was repeated at P2, P3 and P4. Plots show the maintenance of surface marker expression over 4 passages.

3.2.5.3 PDXs maintain key genetic features with passage

FISH was employed to investigate the presence of the MYC translocation in tumour cells after multiple passages through mice. This important feature was maintained in the five PDXs between passage 1 and passage 5 (Figure 3.11).

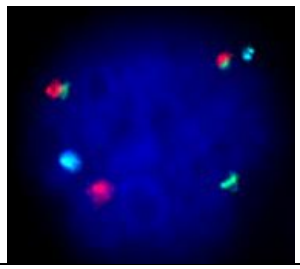
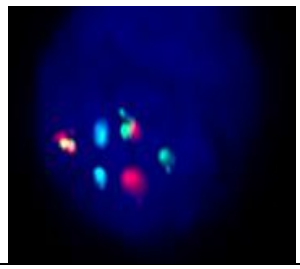
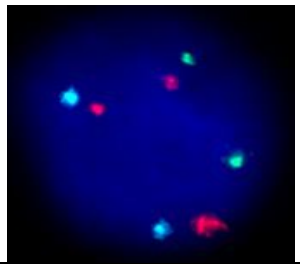
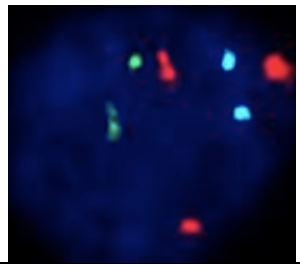
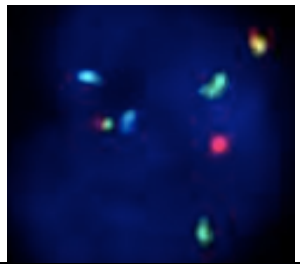
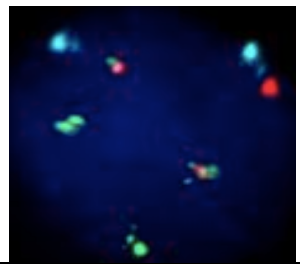
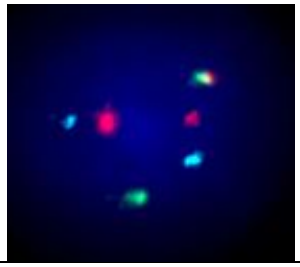
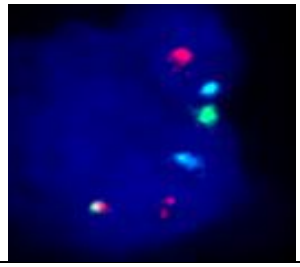
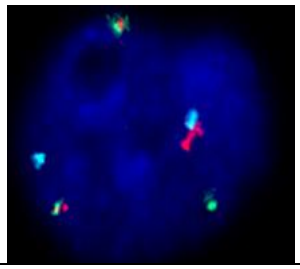
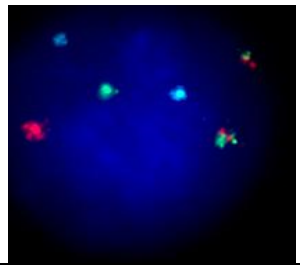
	Signal at P1	% at P1	Signal at P5	% at P5
PDX1	1R1G2F2A 	95	1R1G2F2A 	99
PDX2	3R2G2A 	94	3R2G2A 	96
PDX3	1R2G2F2A 	97	1R2G2F2A 	95
PDX4	2R1G1F2A 	95	2R1G1F2A 	94
PDX5	1R1G2F2A 	100	1R1G2F2A 	96

Figure 3.11. MYC translocation is maintained after five passages in mice

Following passage 1 and 5 in mice, tumours were disaggregated and the presence of the MYC translocation signal investigated by FISH. Images represent one DAPI stained nucleus of a cell. 100 cells were counted per sample and reported. Aqua (A) signal = centromere of chromosome 8, red (R) signal = MYC, green (G) signal = IgH, F = fusion.

However, tumour cells are genetically distinct with sub-clonal variation and so gene expression profiling for 20 genes on RNA isolated from tumour cells after passages 1 and 3 was conducted (Figure 3.12). These 20 genes were chosen based on studies of expression profiles of BL and also included genes associated with stemness and genes identified as important in other PDX studies[58,395]. A close correlation was detected whereby the correlation coefficient between P1 and P3 for each PDX was greater than 0.94, highlighting consistent expression levels of this limited cohort of genes in both passages.

Gene
GAPDH
MYC
BCL2
BCL6
GNA13
ARID1A
ID3
TCF3
ALDH1A1
VEGFA
PIM1
CDK4
CDC7
VEGFA
TERT
SOX2
PTEN
COL1A1
PAX5
MTOR

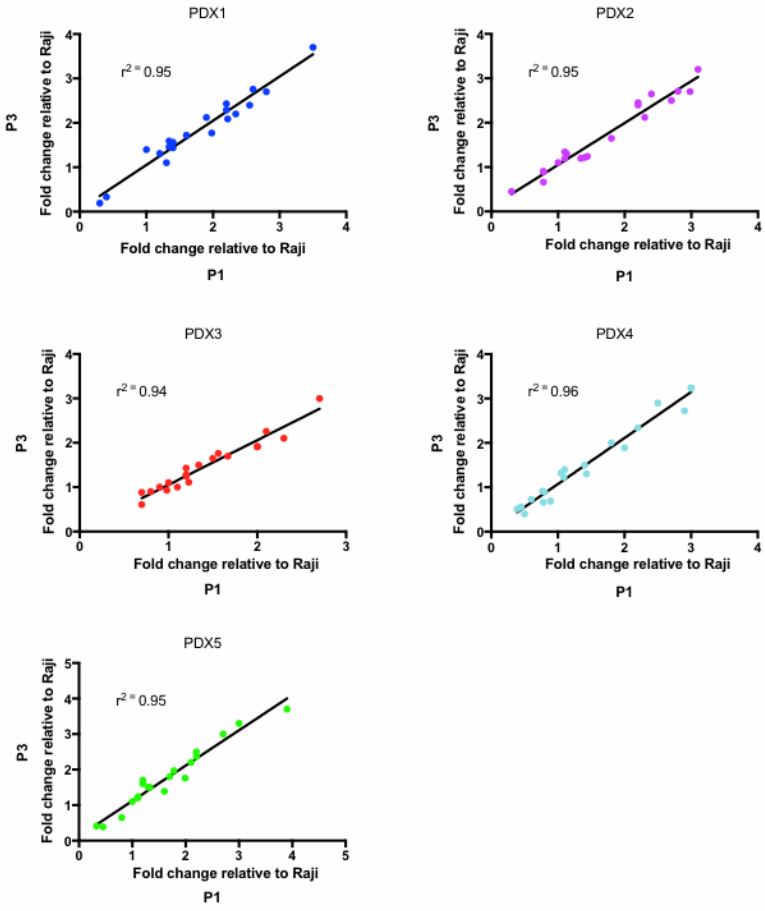


Figure 3.12. Expression levels of 20 genes were maintained in PDXs between passage 1 and 3.

RNA was extracted from PDXs after passage 1 (P1) and 3 (P3). RT-qPCR was carried out on both samples using the same reference cDNA from the Raji cell line as the control. The expression fold change for each gene was determined by $2^{-\Delta\Delta Ct}$.

Taken together, comprehensive phenotyping by histology, FISH, flow cytometry and gene expression highlighted the successful maintenance of tumour properties in the PDX models. Each of these 5 successfully engrafted tumours retain their MYC translocation through passage as well as their histology and gene expression profile for a sample of genes. Importantly, analysis of cell surface protein expression highlighted heterogeneity at both the inter- and intra-tumour level (Figure 3.10). Each tumour faithfully maintains its cell surface protein expression profile over passage, highlighting the existence of an intricate population equilibrium within individual samples. Comparison with the starting tumour population was not made as, for many cases, there was insufficient material to conduct detailed flow cytometry and in addition, the starting tumour population was not pure, being mixed with other cellular components of the original biopsy.

3.3 Discussion

We have successfully established PDX models of paediatric B-NHL and demonstrated their ability to faithfully maintain molecular and phenotypic characteristics through multiple passages. This ability to maintain tumoural heterogeneity ensures the model possesses a high level of clinical relevance and can be used for identifying populations with chemo-resistant potential as well as to investigate targeted therapies.

The utilisation of cell lines has been shown to be lacking in terms of their ability to translate results to the clinic, mainly due to their clonal nature which hinders the study of intra-tumour heterogeneity as well as inter-patient heterogeneity. As survival rates in pB-NHL approach 90% but salvage rates in relapse remain at <30%, it is necessary for us to develop a greater understanding of the different populations within tumours to not only try to reduce the use of toxic chemotherapy but also to develop targeted therapies for the difficult to treat chemo-resistant relapse and refractory cases. The well-established Paediatric Preclinical Testing Program supported by the National Cancer Institute established 51 solid tumours (including sarcomas, neuroblastoma, brain tumours, osteosarcoma) and 10 ALL models and has shown that prediction of response in the clinic is significantly higher when a PDX model is used as opposed to cell line xenografts[19]. Due to the genetic heterogeneity of B-NHL, several different pathways leading to drug resistance have been identified in the adult disease and recent data highlights that this is likely to be the same in paediatric cases[62]. Successful treatment of chemo-resistant pB-NHL may therefore require the rational design of combinatorial drugs targeting multiple populations and pathways specific to different subtypes of B-NHL as well as the development of personalized approaches to address patient-to-patient genetic heterogeneity.

One of the initial considerations in the establishment of a PDX model is the genotype of immunodeficient mouse to use. Several types of immunodeficient mice can be used to establish xenograft models: athymic nude mice, SCID, NOD-SCID, and recombination-activating gene 2 (Rag2)-knockout mice. These strains have been shown to be appropriate for establishment of B-NHL cell line xenografts[396], although they are often not sufficiently immunodeficient for the establishment of PDXs from primary cancer

cells due to their functional innate immune system[397]. With this in mind and with personal knowledge that other researchers have had difficulty establishing B-NHL PDXs we chose to use the NSG model which has enhanced immunodeficiency through a lack of NK cells and has been shown to increase the take rate of numerous human cancers in PDX models[398]. An engraftment rate of 35.7% was observed when considering all specimens with tumour involvement received, increasing to 50% if just considering individual patients. These rates are comparable to many other similar studies with one review of existing data highlighting that engraftment rates typically vary between 23% and 75% depending on the tumour type[399]. Pleural effusion samples had the highest engraftment success rate although PDXs also developed from peripheral blood and bone marrow via the IP and SC routes. In contrast, no cells engrafted via the IV route, regardless of the biopsy origin site. Possible explanations for the failure of the IV engraftment route include the absence of matrigel as well as the less protective niche the IV route might offer. NSG mice have no mature B or T cells or NK cells although leakiness in this phenotype is possible. Furthermore, they contain defective dendritic cells and so tumour cells may be better protected from any potential residual immune functions of the animal when injected into the IP or SC cavities.

Despite the importance of an absent immune system to enable tumour engraftment, this is also one of the limitations of the NSG mouse model given the emerging evidence in recent years of the important role the TME plays in cancer. This area has not been well-studied in BL with traditional views that there is not a strong dependence on the microenvironment for stimuli in BL due to the fact that malignant cells harbour a MYC translocation and cooperating mutations that confer strong cell-autonomous survival and proliferation signals[400]. However, more recent data has shed light on the potentially important role of the TME in mature B-NHL and in DLBCL the importance of non-malignant cells, particularly macrophages and dendritic cells has been reported[401]. Apart from immune cells the TME is made up of stromal cells including resident fibroblast-like cells, endothelial cells and extracellular matrix. Recently the role of stromal components of the TME in mature B cell malignancies has been highlighted and in particular they have been shown to regulate sustained MYC activation facilitating lymphoma cell growth and proliferation[402,232]. Upon adhesion to stromal cells, malignant cells as well

as stromal cells produce and secrete a host of soluble factors associated with malignant B cell growth[403–405]. Medina *et al*. showed that B-cell lymphoma adhesion-induced BAFF expression by bone marrow stromal cells promoted cell growth in B cell lymphomas via activation of NF-κB signalling[406]. IL-6 which is another prototypical growth and survival factor in B cell lymphomas[407] is also produced by the TME and initiates ERK1/2, PI3K and JAK/STAT3 signaling[408]. Numerous other cytokines have also been shown to contribute to the cross talk between stromal and lymphoma cells including IL-7, IL-4, IL-8, VEGF, CXCR4, MIP-1, Jagged-1, and insulin growth factor (IGF)-1[409–413].

Importantly, emerging evidence has shown that TME also has a critical role in supporting stemness properties in B-NHL. The renewal and differentiation capacity of germinal centre lymphocytes upon antigen stimulation has suggested that secondary lymphoid organs are ideal niches to confer lymphocyte stemness features through stroma-mediated genetic and epigenetic means[402]. Given the close contact of lymphoma cells and stromal cells in the lymph node and bone marrow microenvironment, malignant reprogramming may be achieved through stroma-mediated regulation of genes controlling a stemness pathway[232]. Recent studies on TME have revealed that stromal cells can modulate stemness by influencing B cell differentiation and enhancing lymphoma cell clonogenicity[232,233].

Due to the emerging evidence of the role of the TME in B-NHL as well as the introduction of Rituximab and potentially the future use of other immune therapies for the treatment of pB-NHL further investigation in this area is required. It will be necessary for PDX models to possess a human immune system component to facilitate the study of immune-cancer cell interactions and preclinical assessment of cancer immune therapies. Several different methods exist to develop humanized mice including the transplantation of total peripheral blood or tumour-infiltrating lymphocytes into immunodeficient mice, or the transplant of CD34-positive human hematopoietic stem cells or precursors, either alone or in combination with additional human immune tissues (e.g. human thymic tissue) into immunodeficient mice[301]. Each method has its complexities and with the high expense these are not in common use currently. However, now that we have established and characterised a PDX resource that will

continue to expand, exploring human immune system-conditioned PDX models is a likely direction to improve pre-clinical studies of pB-NHL.

The most important consideration when utilising PDXs as a model for cancer is how well it represents the original patient disease. Due to insufficient biopsy material we used passage 1 compared to later passage tissue to monitor the effect of serial passage on the tumour cell characteristics at the histological, gene expression and surface protein level. However a number of lines of evidence in our study also highlight that it is likely changes did not occur to tumour properties between biopsy and passage 1. All tumours were almost 100% MYC translocation positive after passage 1 and in the two cases where the exact initial translocation signal was available from the diagnostic report, this signal was maintained after passage 1. Furthermore patient sample 2, which produced PDX1 from a pleural effusion, showed that surface marker expression in the biopsy sample MNC population was maintained after passage 1 and in patient sample 3 peripheral blood sample (PDX2), any cell populations that were lost from the biopsy sample were shown not to be $CD20^+$ cells and hence most likely healthy immune cells that fell into the same forward and side scatter gate. Of course, it is not possible to categorically conclude this from our data so it is also possible they may have been tumour cells that did not adapt to the *in vivo* environment. Several studies in the literature also support the maintenance of properties between biopsy and passage 1 including a study of MCL PDXs which sequenced 1,212 cancer-associated genes and found no mutations introduced after passage 1[392]. Furthermore in a study of adult DLBCL PDXs, a comparison of the mutant allele fraction in primary specimens and the associated PDXs indicated that these models retained the complex genetic signature of primary samples[385].

Analysis of tumour properties between subsequent passages from P1 onwards in our model was also an encouraging indicator of the ability of this model to maintain the biological and molecular heterogeneity of the tumours. Nearly every cell at passage 5 was MYC translocation positive with maintenance of the exact signal from passage 1, highlighting that tumour cells are not losing this key BL characteristic. Furthermore, the correlation coefficient for gene expression levels between P1 and P3 was greater than

0.94 for all 5 PDXs. However, we acknowledge that this was a limited set of genes and going forward to further validate the PDXs we would examine expression of a broader array of genes as well as carry out SNP arrays or whole exome sequencing to get an overview of the mutational landscape of the tumours through passage. At the cellular level, the histology of each tumour was maintained over five passages and significantly the heterogenous surface protein expression profile was also maintained highlighting the potential importance of a tightly controlled population equilibrium within the tumours.

Overall, the first well-characterised PDX resource for pBL has been established showing that, at least for some of the PDXs, initial tumour heterogeneity is maintained between the biopsy and the first passage. Furthermore, the characteristics of the tumours in terms of cell surface protein expression, morphology and the presence of the driving MYC translocation are maintained through multiple passages in the mice and provide us with a valuable resource to investigate the functional and clinical significance of the heterogenous populations in these BL tumours. PDXs offer a greatly improved route towards personalised medicine. A notable example is a pilot clinical study in which pancreatic PDX models were used to guide treatment for 11 patients with advanced cancers. 17 treatment plans were devised with 15 of these resulting in durable partial remissions[414]. This type of modelling is beneficial in cases where few established treatments exist such as in the case of relapsed of refractory pB-NHL. The existence of a PDX that represents the full heterogenous nature of the patient disease would allow for the development of targeted therapy plans, focusing on the correct population of cells.

Chapter 4. Heterogenous populations exist in paediatric B-NHL including a side population

4.1 Introduction

With evidence generated thus far highlighting that the PDX model accurately maintains tumour properties through passage, the heterogeneity of tumour cell populations within these PDX tumours was investigated utilising limiting dilution assays, cell surface phenotyping and functional assays. The fact that malignancies, both epithelial-derived and haematological, show morphological and physical heterogeneity has been known for many years, since the first pathological assessments were conducted[204]. However, more recently this knowledge has been expanded upon by increasing awareness of variation in both molecular and functional biology, as assessed by *in vitro* and *in vivo* assays. These developments have led to the search for populations within heterogenous malignancies that are capable of maintaining disease and most importantly are responsible for clinical relapse.

The stochastic model of cancer was proposed by Nowell *et al.* in 1976 and utilises the concept of clonal evolution to explain heterogenous populations present in tumours[192]. This model proposes that each cell in the whole tumour population has the ability to produce a malignancy. Mutations arising in tumour cells confer a selective growth advantage and hence lead to the observed heterogeneity. Both endogenous and exogenous factors can result in this selective pressure. Endogenous effects include variation in growth, apoptosis, metabolism, and other "hallmarks of cancer." Non-cell autonomous factors causing selective pressure include cytokine concentrations, cell–cell interactions and the niche TME[206].

Alternatively, the CSC model suggests a hierarchical organization of cells, whereby a small subset has the ability to sustain tumorigenesis and generate heterogeneity through differentiation[415]. The term CSC is generalized to define a distinct population of cells with stem cell-like properties and not necessarily a transformed tissue-specific stem cell. These cells have self-renewal properties and can undergo differentiation to give rise to a transient-amplifying (TA) cell population with limited replicative potential.

The TA cells are responsible for forming the bulk of the tumour[416,417]. CSCs by definition must be able to give rise to tumours upon injection into immunocompromised mice and the tumours they produce should be serially transplantable as a demonstration of long-term repopulation capacity[202].

Until recently these two models, i.e. stochastic versus CSC, were considered mutually exclusive. However recent advances in knowledge have highlighted the blurred lines in this area of study and it is proposed that the two models of cancer heterogeneity can be harmonized[202]. Greater knowledge of cancer genetics, epigenetics and the tumour microenvironment has highlighted the important role each of these has in defining stemness and less focus is placed on tumours strictly fitting either model of heterogeneity. This theory compliments the findings discussed in a review by Visvader in which CSCs were described to have variable phenotype, to be present as multiple clones within individual tumours and to undergo genetic evolution[415]. Finally, it is not incomprehensible that non-CSCs may have the potential to transition to CSC-like cells[415].

The CSC model has become more accepted and its therapeutic implications realised, although its study has been hampered both by a lack of consistency in the terms used for these cells and by how they are defined. A study of the literature has identified a lack of clarity in the use of the terms CSC, tumour-initiating cell (TIC) and tumour-propagating cell (TPC). Reviews by Visvader and Lindeman as well as Kreso and Dick have addressed this and I will utilise their definitions throughout this thesis[415,202]. Contributing to a lack of clarity is the fact that not all tumours harbour CSCs. For example, in melanoma, the high proportion of tumorigenic cells (as many as 50%) assayed under more permissive conditions and with a wide spectrum of markers (e.g., CD271), argues against a CSC model of heterogeneity[212]. Cells that form tumours in this case cannot be deemed CSCs but are more appropriately called TPCs. Conversely, the term TIC has been reserved for cells with characteristics closer to CSCs. In the case of primary colorectal cancers, three different cell types were all assigned the title of TICs on the basis of clonal sphere cultures derived from individual patient tumours[418]. These included a rare subset of CSC-like cells that maintained tumour growth on serial transplantation, a TIC with limited self-renewal capacity (therefore not defined as a

CSC), and a more latent CSC that was activated only in secondary or tertiary transplantation assays. Dick bases his definition of a CSC on the defining characteristics laid out by Clarke *et al.* in 2006 and Nguyen *et al.* in 2012, whereby CSCs represent a distinct population that can be prospectively isolated from the remainder of the tumour cells and can be shown to have clonal long-term repopulating and self-renewal capacity[194,200]. Because prospective identification and isolation is difficult, they and others refer to populations which cannot be prospectively isolated but have CSC properties as TIC or leukaemia-initiating cells. Hence, TICs are defined by their ability to (1) generate a xenograft that is representative of the parent tumour, (2) self-renew as demonstrated by serial passage in a xenograft assay at clonal cell doses, and (3) give rise to daughter cells that may possess proliferative capacity but are unable to establish or maintain the tumour clone upon serial passage. Their only difference compared to a CSC is that they cannot be prospectively isolated[202].

The study of heterogeneous tumour cell populations and in particular the identification of CSCs is a difficult task and researchers have used a range of techniques over the years. Xenotransplantation has been widely used for both the identification and quantification of cells capable of propagating tumour growth. As mentioned earlier, the importance of identifying these cells in order to be able to separate them and class them as CSCs is important and hence flow cytometry and FACS became powerful tools in this area. Early research focussed on distinguishing heterogenous populations by examining cells for the presence of cell surface proteins typically associated with differentiation status on the assumption that CSC might express proteins associated with stemness[419]. FACS was used in the seminal studies published by Lapidot *et al*. in the late 90s when CSCs in AML were identified by expression of CD34, a cell surface protein associated with HSCs[197]. Over time these studies were expanded to a whole host of proteins, not just those associated with tissue-specific stem cells. Indeed, several of the proteins present on varying percentages of tumour cells in our PDXs have previously been implicated in CSC studies including CD24 in breast, gastric and pancreatic cancer[420,421], CD9 in ALL[422] and CD44 in a large array of cancers[421,423,424].

The process of identifying CSCs by cell surface phenotyping can be challenging as a large number of surface proteins are often expressed and the CSC compartment could be a population identified by expression of a combination of any number of these proteins. For this reason, I focussed on identifying CSC-like cells based on their ability to fulfil functional criteria. One such assay is the Side Population (SP) assay; this technique developed by Margaret Goodel identifies a fraction of cells with the ability to efflux the Hoechst dye[334]. These cells were later shown to express ABC transporters likely evolutionarily conserved on cells with stemness properties as a protection mechanism[333].

Aims of this chapter

- Conduct limiting dilution assays with PDXs to estimate the proportion of cells within the bulk tumour capable of propagating tumours
- Investigate whether heterogeneous cell surface protein expression within the PDXs can be employed to identify cells with CSC properties
- Determine the optimal conditions for the detection of a SP in these PDXs
- Investigate the potential of the SP to meet the characteristics of a CSC population
- Attempt to phenotype the SP by examining them for the expression of defining cell surface proteins.

4.2 Results

4.2.1 Limiting Dilution Assays show the frequency of tumour propagating cells to range from 1 in 159, to 1 in 1120 in different PDXs

Traditionally, the existence of a distinct tumour propagating population has been demonstrated through limiting dilution assays whereby the lowest number of cells from which it is possible to generate a tumour is indicative of the percentage of TPCs within a bulk tumour. This technique is not without its caveats, but was employed in this study as a crude method to identify heterogeneous populations of tumour cells, particularly given the presumed homogeneity of BL. Quantifying the frequency of cells with tumour propagating ability is useful as it can give an estimation of whether all cells within the tumour can produce further tumours or if there is a distinct population alone with this ability. A range of tumour cell numbers from PDX1-3 inclusive were injected into cohorts of mice and Poisson statistics used to calculate the frequency of cells capable of initiating xenografts[425,426]. Cells were isolated from a tumour grown in an NSG mouse after passage 2, disaggregated and re-injected suspended 1:1 in Matrigel, in limiting dilutions into NSG mice. When as few as 25 cells from PDX1 were injected into recipient mice, 1/3 developed a tumour although when 100 cells were injected into each of three mice no tumours formed (Table 4.1). Following ELDA (Extreme Limiting Dilution Analysis)[427] of the experimental data using a confidence interval of 95%, the predicted TPC frequency for PDX1 was determined to be 1/285 (range = 1/79 to 1/1032). Identical analysis of PDX2 and PDX3 tumours yielded predicted TPC frequencies of 1/1120 and 1/159 respectively. Interestingly, although this is a broad estimate, it does highlight an important observation – not all cells in any of the PDXs are capable of propagating tumours, in fact quite a small proportion of cells seem to have tumour propagating potential.

No. of cells injected	No. of mice injected	No. of mice developing tumours
PDX1		
1000000	1	1
100000	2	2
10000	2	2
1000	3	3
100	3	0
25	3	1
5	3	0
PDX2		
1000000	1	1
100000	2	2
10000	2	2
1000	3	2
100	3	0
25	3	0
5	3	0
PDX3		
1000000	1	1
100000	2	2
10000	2	2
1000	3	3
100	3	1
25	3	1
5	3	0

Confidence intervals for 1/(TPC frequency)		
Lower	Estimate	Upper
1032	285	79.2

Confidence intervals for 1/(TPC frequency)		
Lower	Estimate	Upper
4456	1120	282

Confidence intervals for 1/(TPC frequency)		
Lower	Estimate	Upper
611	159	41.5

Table 4.1. Limiting Dilution Assays using ELDA analysis estimate the frequency of tumour propagating cells in PDX1, 2 and 3.

4.2.2 Cell surface protein expression profiling indicates tumour cells with varying levels of tumourigenic potential

Flow cytometry is a powerful technique that has been utilised to identify heterogeneity within tumours by examining the cell surface expression of a whole host of proteins indicative of lymphoid cell developmental stages and functions; for example, in humans $CD34^{+}CD38^{-}$ is the immunophenotype identifying HSCs with other proteins including CD90 (Thy-1) and IL7Rα refining this phenotype further[428]. Other proteins are indicative of later developmental stages such as CD21 and CD24 expressed on immature B cells, while CD27 and CD40 are expressed on activated germinal centre B cells. The variable expression levels of these and many other cell surface proteins within the PDX tumours is displayed in Figure 3.10 in chapter 3 and provided the basis for experiments to assess the functional properties of these distinct cell compartments, specifically their ability to seed new tumours.

Initial experiments were conducted using paediatric BL cell line xenografts as opposed to patient-derived tumours. In particular, Ramos cells which have a broad array of proteins expressed on sub-populations of cells including CD9, CD40, CD44 and CD49D were investigated. Cells were sorted into positive and negative fractions for each specific protein by FACS. Positive and negative cells ($5X10^{5}$) were then injected into separate NSG mice in triplicate and tumour growth monitored until the maximum size limit was reached (12mm in any one dimension). The tumours were then resected and the size verified. All cellular fractions produced tumours that reached the size limit within a maximum of 41 days (Figure 4.1). However, there were significant differences in the tumour forming capacity of the CD9, CD44 and CD49D-expressing cell populations; the $CD9^{-}$, $CD44^{+}$ and $CD49D^{+}$ fractions reached the tumour size limit significantly faster than their opposite cell populations (i.e. $CD9^{+}$, $CD44^{-}$ or $CD49D^{-}$). This highlights that each of these populations, positive or negative for each surface protein, can be termed TPCs although it does lend support to the concept of multiple subclones existing within the cell line that possess different levels of tumourigenic potential.

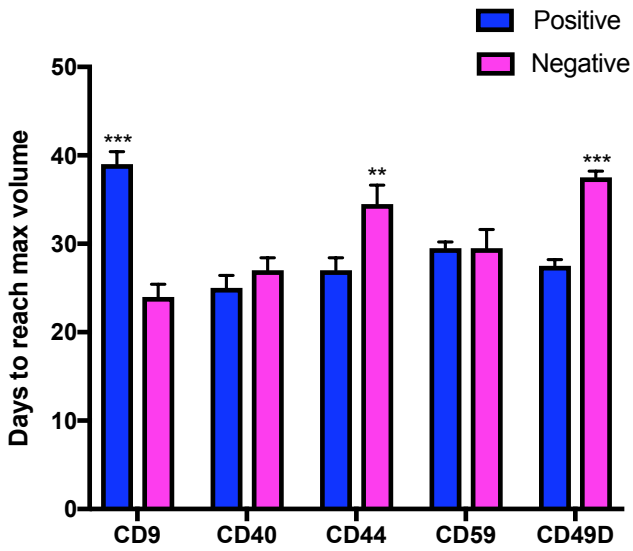


Figure 4.1. Differential tumorigenic potential of sorted cell fractions in Ramos cell line xenografts.

*Cells were stained with antibody and sorted by FACS into positive and negative populations before 5X10^5 cells were injected with matrigel into 3 NSG mice for each cell population. The time taken for the tumours to reach the maximum allowed size was determined. * p<0.05, **p<0.005, ***p<0.001*

Once experiments had been conducted and optimised with the cell lines, and given the particularly striking significance associated with CD9, positive and negative populations were sorted from each of the four PDXs that also contained CD9-expressing subclones of cells (Table 4.2). This was conducted with PDX tumours isolated after passage 2 and cells injected into two further mice for each cell population. The tumours were measured on a daily basis and the time to produce maximum allowable tumour size recorded. Interestingly, the CD9 negative population again had significantly increased growth potential compared to the CD9 positive population in all 4 of the PDXs (Figure 4.2).

PDX #	CD9+ Cells (%)	CD9- Cells (%)
2	91	9
3	98.2	1.8
4	71	29
5	87.3	12.7

Table 4.2. Proportion of CD9 positive and negative cells in 4 PDXs (PDX1 is CD9 negative)

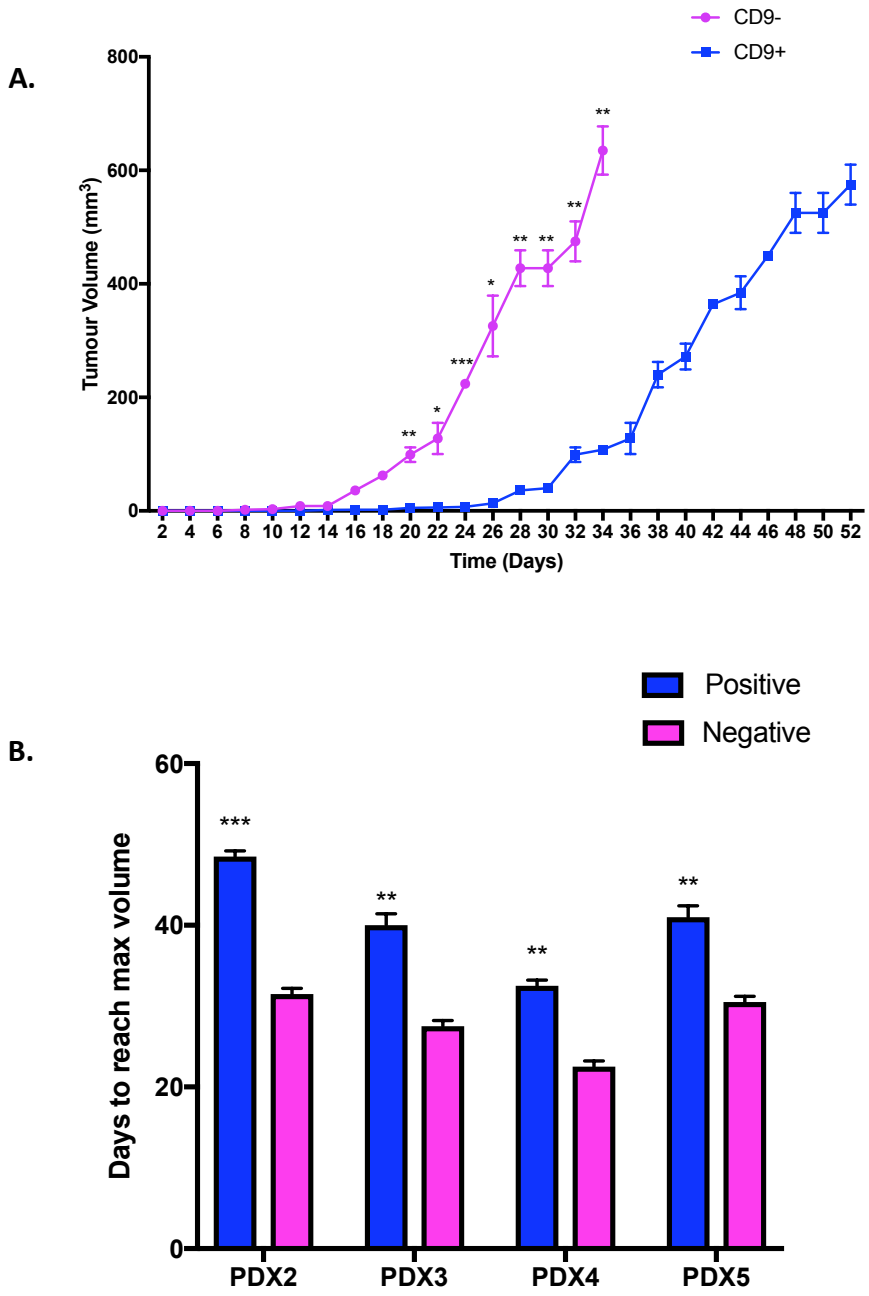


Figure 4.2. Differential tumorigenic potential of sorted CD9 positive and negative fractions isolated from PDX tumours.

*Cells were labelled with antibody and sorted by FACS into CD9 positive and negative populations before 5X10^5 cells were injected with matrigel into two NSG mice each. The time taken for the tumours to reach maximum allowed size was measured. (A) Representative data for PDX2 highlights the significant difference in growth rate at each time point. (B) Data show the times taken for tumours to reach the maximum size, comparing populations either positive or negative for CD9. *p<0.05, **p<0.005, ***p<0.001*

CD9 expression is just one example of a surface phenotype that resulted in differential growth rates amongst tumour populations. Further examples of increased tumourigenic potential based on cell surface protein expression are presented in Table 4.3. Both the intra-tumoural and inter-tumoural heterogeneity of these cell surface proteins within our PDX resource highlight that the landscape of pBL is likely more complex than previously thought.

PDX	Fraction with greater tumourigenic potential	Proportion of cells (%)	*p*-value
1	CD44+	89	0.008
2	CD24+	86	0.009
3	CD49D+	15	0.007
4	CD90+	88	0.0009

Table 4.3. Cell surface protein expression identifies tumour cell populations with significantly increased tumourigenic potential *in vivo*.

Tumour cells of each cell fraction, positive or negative for the indicated proteins, were injected into two NSG mice and the time taken to reach a size of 12mm in any direction was recorded.

4.2.3 Cells expressing primitive surface proteins are present in PDX tumours

As well as examining the cell surface proteins that characterise BL and those identified in other CSC studies, a broad array of antibodies were utilised to identify populations of cells expressing markers associated with HSCs and 'stemness'. In particular, the primitive cell surface proteins CD34, CD90, CD117 and ABCG2 were investigated. CD34 is a HSC marker and identification of cells expressing this protein may indicate a CSC population that originated from an early progenitor cell. Like CD34, CD90 is a protein expressed on pluripotent HSCs[429]. Cell surface proteins associated with early stem and progenitor cells were identified in each of the PDXs. Of particular note is CD90 positivity for almost all cells in PDX4 (Figure 4.3). The expression of proteins associated with stemness on tumour cells with a presumed mature B cell origin is an interesting observation and these data suggest that further characteristics of stemness in the PDXs should be explored. Representative graphs are shown for PDX4 (Figure 4.3) with identified populations frequencies for other PDXs shown in Table 4.4.

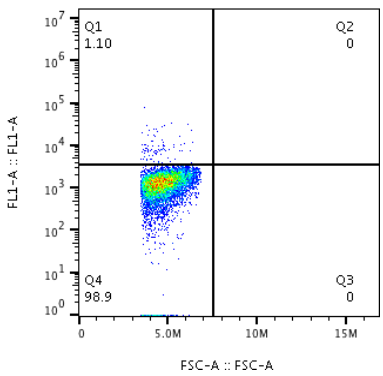
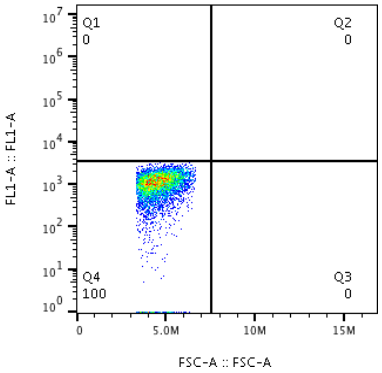
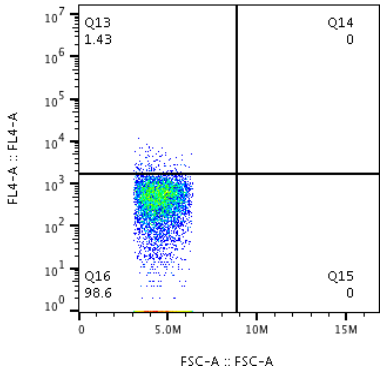
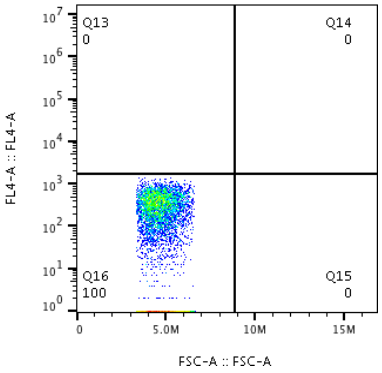
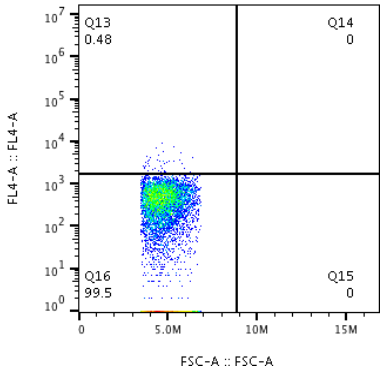
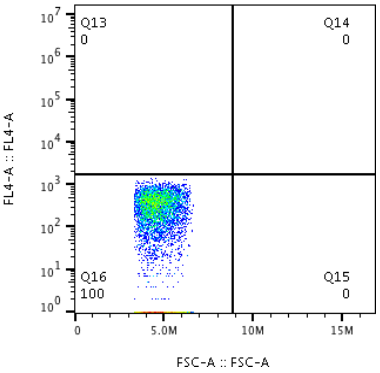
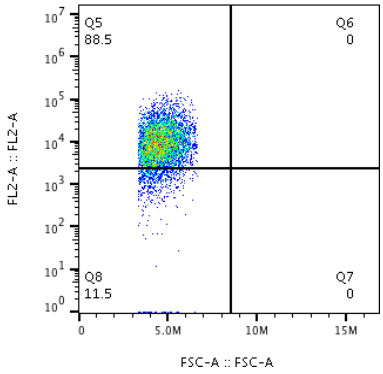
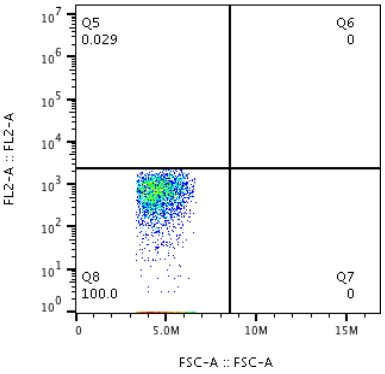
	Conjugated antibody	Isotype Control
CD34		
ABCG2		
CD117		
CD90		

Figure 4.3. Stemness associated cell surface proteins including CD34, CD90, CD117 and ABCG2 are expressed on PDX4

Cells from PDX4 tumours were labelled with stem cell associated surface proteins. Gates were assigned using unstained negative controls and the absence of non-specific binding verified with appropriate isotype controls. Plots are representative of two biological replicates.

PDX	% CD34+	% ABCG2+	% CD90+	% CD117+
1	1.34	1.13	0.7	0
2	0.84	1.09	0	0
3	0.42	0.82	0	0.35
4	1.10	1.43	88.5	0.48
5	0.67	1.35	1.13	0

Table 4.4. Proportions of cells with each stemness associated surface protein marker in five PDXs

4.2.4 A side population (SP) is present in all 5 PDXs

Based on the difficulty of prospectively identifying a CSC population by cell surface phenotyping alone, the SP functional assay was used to determine the presence of cells with stemness features based on their ability to efflux the Hoechst dye. It has been noted that in order to identify the SP, technical parameters such as cell density, dye concentration, time of incubation, data analysis and correct gating must first be carefully optimised[356]. Optimisation was initially carried out using PDX-derived tumour cells harvested from mice at passage 1 or 2 when the tumour was at maximal size. In the first stage, incubation time and concentration of Hoechst dye was optimised. The presence of a SP cells was determined by examining their Hoechst 33342 fluorescence intensity (Figure 4.4).

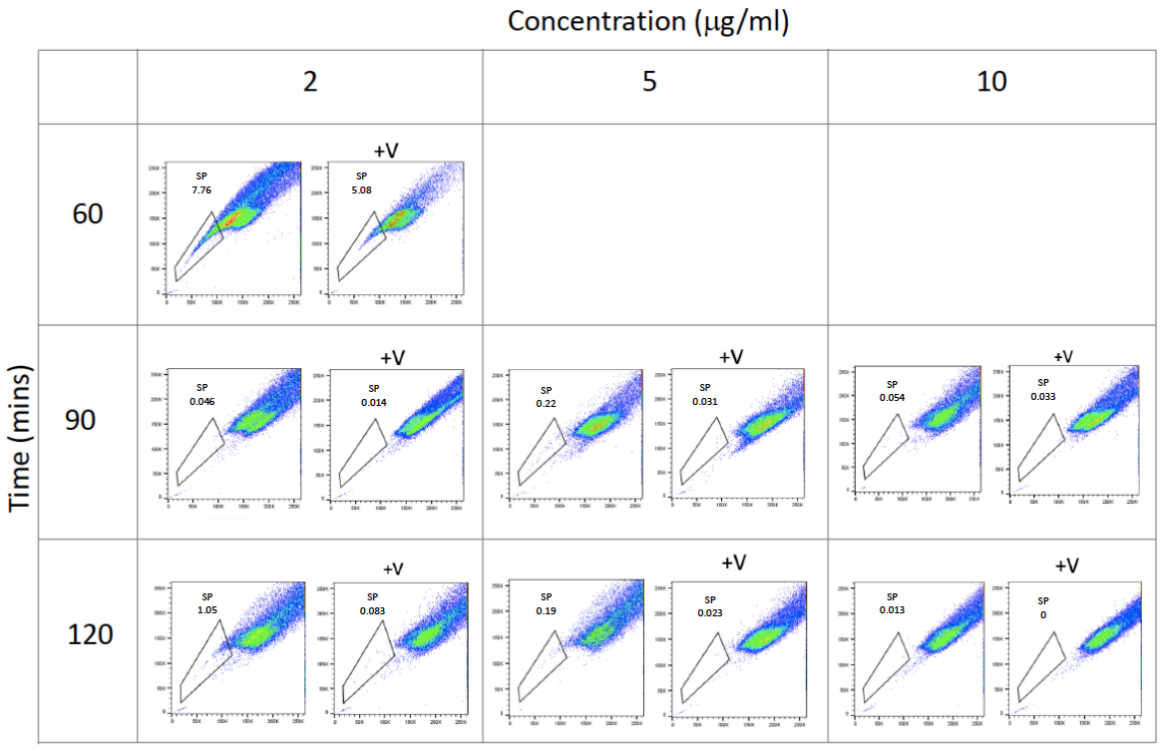


Figure 4.4. Optimisation of the protocol for the detection of SP cells in PDX1

The SP assay was carried out at the indicated times and Hoechst dye concentrations. Data are representative of technical triplicates. A Verapamil control (+V) was included for each condition.

Following 60 minutes incubation with 2µg/ml Hoechst dye, there is a large SP of 7.7% which is not significantly diminished on verapamil co-incubation, implying that this is not a true SP and rather is likely due to the short incubation time whereby not all cells had sufficient time to take up the Hoechst dye. Other combinations of dye concentration/incubation time including 5µg/ml for 90 minutes and 5µg/ml at 120 minutes led to the detection of small SP fractions. However, the largest SP (1.05%) that was diminished following co-incubation with verapamil was seen with 2µg/ml dye incubated for 120 mins. Optimisation was also carried out for the other four PDX tumours, all showing similar results, except PDX4 which required a stronger

concentration of Hoechst dye and incubation for 120 mins to identify the SP. The optimal conditions for detection of a SP in each PDX are shown in Table 2.4 of chapter 2. Figure 4.5 shows the SP present in each of the five PDX-derived tumours at passage 2.

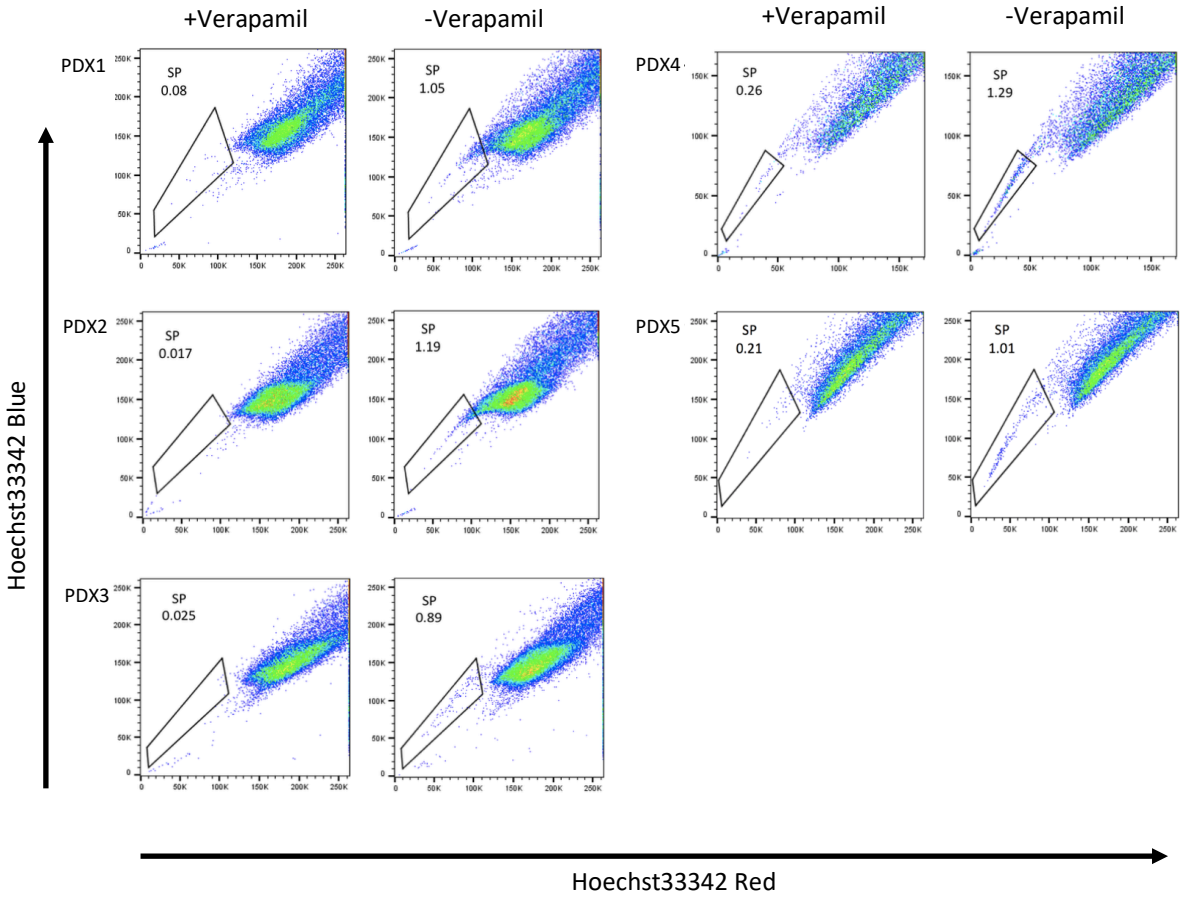


Figure 4.5. A Side Population is detectable in tumours derived from all five PDX models

Following optimisation of the staining protocol for all 5 PDXs, cells were incubated with Hoechst dye at the optimal concentration and time, with and without verapamil. Data are representative of three biological repeats.

At times the exact positioning of the SP gate was difficult to determine, especially if a small number of cells remained in this gate in the verapamil control. Viewing the plot in density mode made it simpler to distinguish the overall populations and hence was often used to assign the SP gate (Figure 4.6).

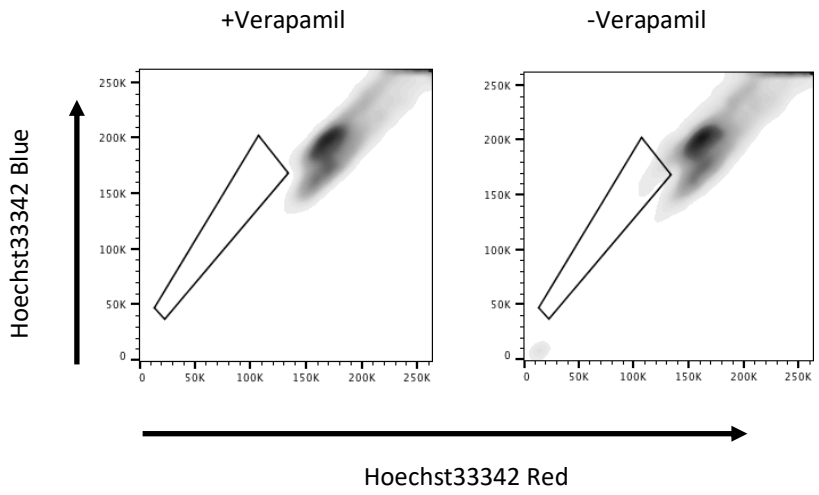


Figure 4.6. A density view of SP plots was used to assign the SP gate

In cases where a small number of assumed SP cells were still present in the verapamil control, density view made it easier to assign the location of the SP gate.

Sufficient material was available from initial biopsy sample of PDX5 to enable SP analysis to be performed on the original sample. Significantly, a SP was identified in the isolated MNCS from this patient pleural effusion sample (Figure 4.7) highlighting that this population is not an artefact of engraftment and, similar to the data in chapter 3, provides strong evidence that the PDX model is an excellent representation of the original patient tumour population.

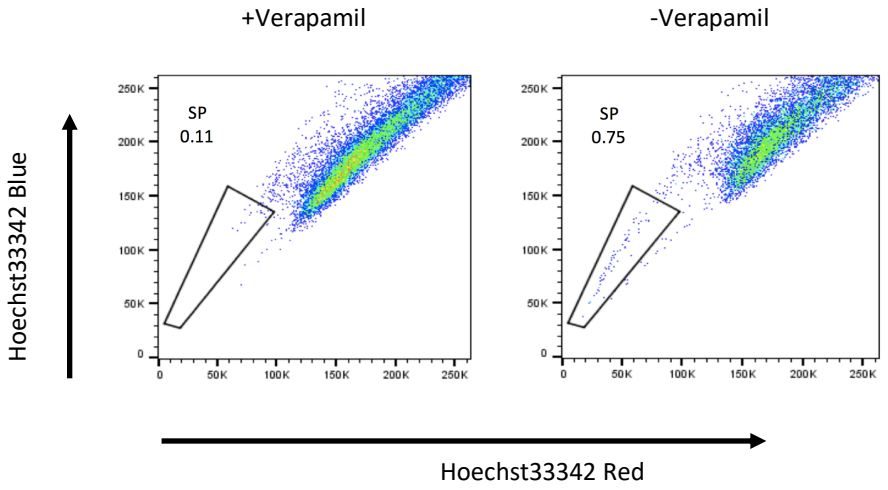


Figure 4.7. A Side Population was present in the biopsy sample from which PDX5 was derived

MNCs were isolated from a pleural effusion biopsy sample and stained with Hoechst dye with and without verapamil control and the presence of a SP was determined by flow cytometry.

4.2.5 The SP fraction of the bulk tumour population cannot be identified by a distinct cell surface protein expression profile

Each of the PDX tumours contain cells expressing cell surface proteins indicative of a primitive origin (Figure 4.3, Table 4.4). The maintenance of these minute populations of cells over time is an interesting observation and suggests they may have an important role in the tumour dynamics. As ABCG2 expression is a defining feature of the SP phenotype, co-expression of this surface protein with all other markers identified within the PDX4 tumour population was investigated to try and refine the SP phenotype. PDX4 shows expression on subpopulations of cells within the whole tumour of CD9, CD24, CD49D, CD34, CD133, CD117 and CD44 as well as strong expression of CD90 (Figure 3.10) and many of these have been associated with characteristics of stemness. Interestingly, although for most surface proteins ABCG2 was expressed within the positive and negative fractions (e.g. CD9), ABCG2^{+} cells were almost completely CD24, CD49D and

CD90 positive (Figure 4.8). Furthermore, $CD133^{+}$ cells were exclusive to the ABCG2 positive fraction. These data suggest that it may be possible to assign a cell surface protein expression profile to the SP present in our PDXs.

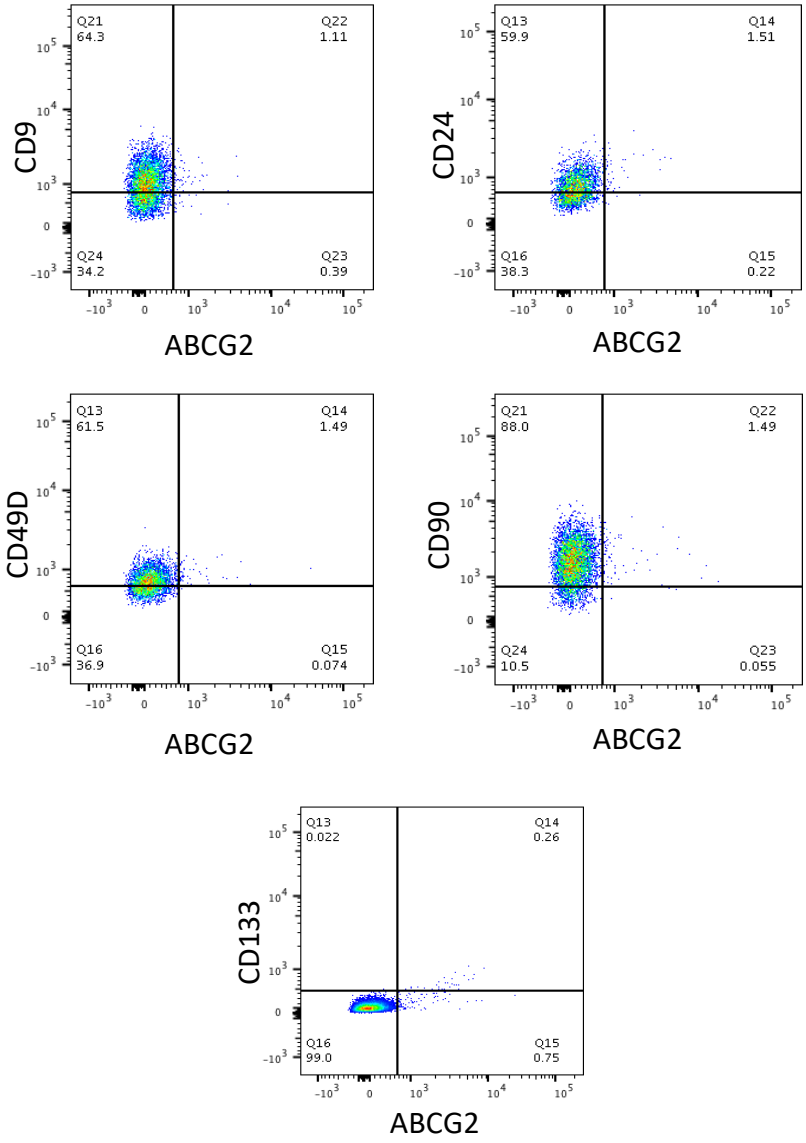


Figure 4.8. ABCG2 positive cells express cell surface proteins associated with cells having properties of stemness

PDX4 cells were incubated with antibodies against ABCG2 and a range of different surface proteins as indicated. Data are representative of three repeats.

However, whilst significantly enriched in the SP compartment of PDX4, cells expressing ABCG2 were not exclusive to the SP (Figure 4.9). Furthermore, proteins shown to be specific to the ABCG2 compartment in PDX4, such as CD133 (Figure 4.8), were also found in both MP and SP fractions (Table 4.5). Regardless, a change in the percentage of SP cells expressing certain proteins was observed, for example, CD9 was expressed by a considerably lower percentage of SP cells in PDX4 (Figure 4.9). Similar results were seen for all five PDXs whereby expression of stemness-related cell surface proteins was increased in the SP fractions (Table 4.5). ABCG2 was enriched in the SP of all 5 PDXs and CD90 which is expressed in PDX 1, 4 and 5 was present on a higher percentage of SP cells, indeed all tumour cells from PDX4 SP are $CD90^{+}$. CD34 positive cells are also increased in the SP of all five PDXs. Apart from proteins traditionally associated with stem cells, other interesting results were noted. For example, CD9 expression is greatly decreased on the SP of four PDXs while CD49D is on a significantly higher fraction of the SP of 3 out of 5 PDXs. CD24 is already highly expressed in four PDXs although in PDX4 the percentage of cells expressing CD24 in the MP is 57%, compared to 83% of SP cells expressing this protein. Finally, CD44 which is only expressed in PDX1 and 2 is also significantly higher in the SP fraction than the MP. Whilst it is not possible to identify any cell surface protein specific to either MP or SP, these data offer an interesting insight into the expression profile of the SP. It is likely that the SP enriches for the CSC population but may not in its entirety constitute the CSC compartment. The CSC could still be identifiable by a specific cell surface protein profile but this may be a combination of multiple proteins and may differ for each tumour.

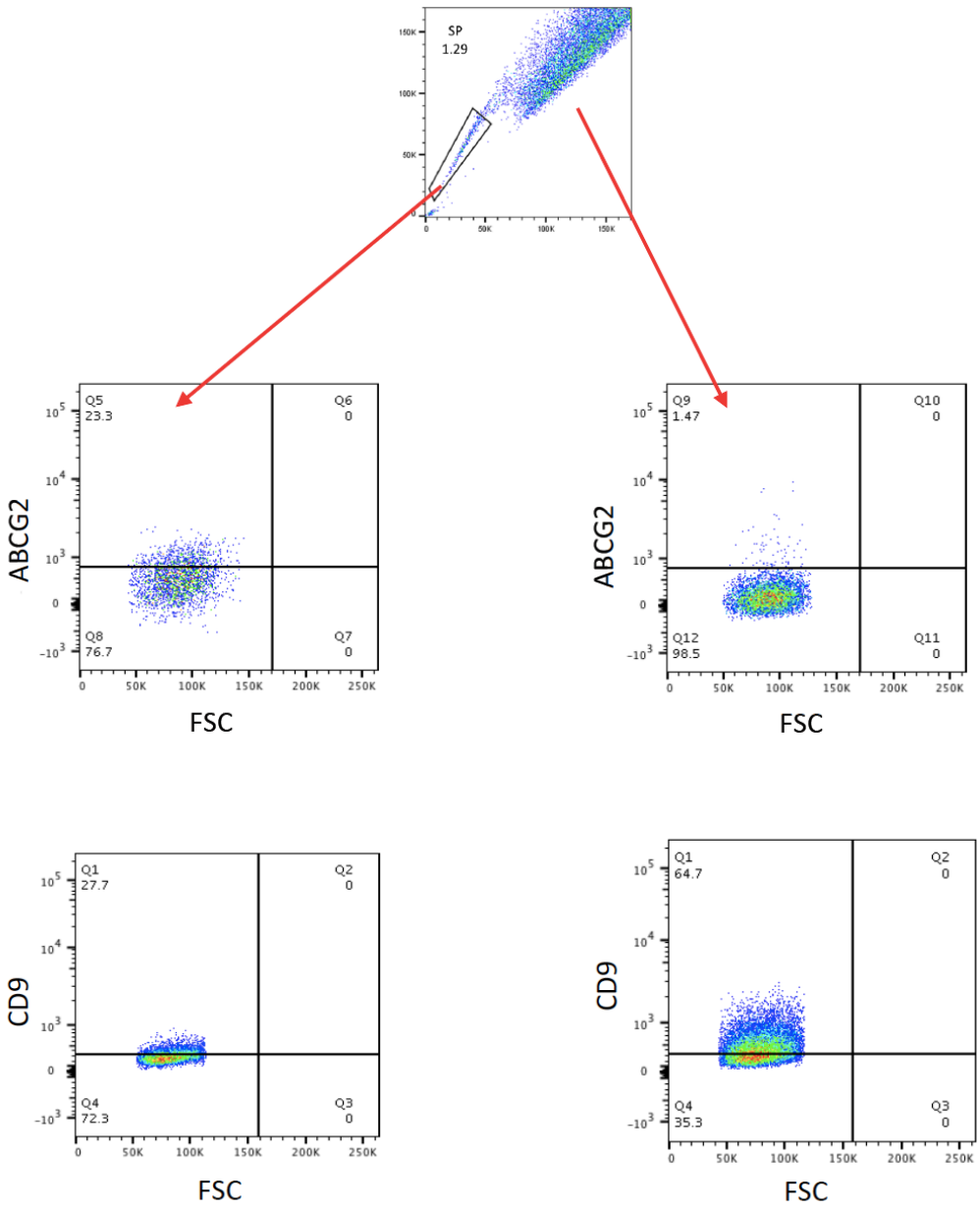


Figure 4.9. ABCG2 is expressed on a higher proportion of SP cells while the CD9 population is decreased in the SP

PDX4 tumour cells were incubated with Hoechst dye and the SP and MP separated by FACS. Both populations were incubated with antibodies against a range of surface proteins and analysed by flow cytometry. Data are representative of two biological replicates.

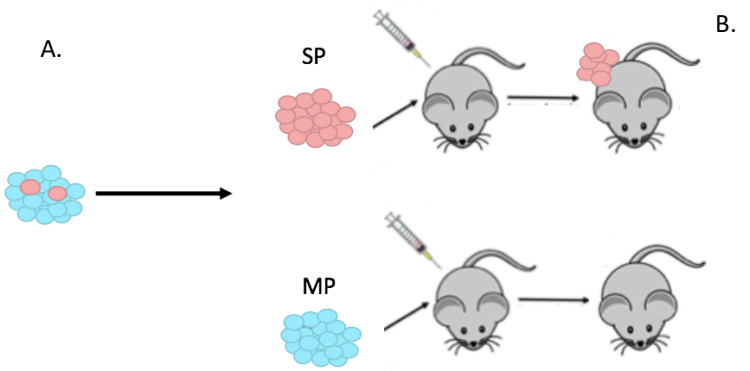
	PDX1		PDX2		PDX3		PDX4		PDX5	
	MP	SP	MP	SP	MP	SP	MP	SP	MP	SP
CD9	0	0	89.4	52.1	99	63.9	64.7	27.7	88	41.2
CD10	99.7	89.3	94	95.2	100	81.6	94	94.8	99.9	96.2
CD19	98	98.3	99.4	99.8	98.9	99	99	100	99.9	100
CD20	98.9	99.1	96.1	95.2	94.4	93.8	97	96.6	96	97.3
CD21	22.4	22.9	0	0	3.2	2.9	2.6	3.1	3.5	4.2
CD24	99.9	100	89	99.6	99	100	57	83.4	100	100
CD27	32.3	29.9	98	99.4	10.2	9.2	99.1	95.6	98.6	99.2
CD34	1.34	3.9	0.84	5.9	0.4	4.8	0.39	4.2	0.5	3.8
CD40	15	18.7	0.6	0.71	4.6	4.3	10.2	9.8	50.6	55.9
CD44	89	97.9	0.5	2.2	0	0	0	0	0	0
CD45	99.7	100	96.7	98.2	100	100	98.5	99.2	99.1	99.4
CD49D	44	76.8	92.4	91.2	15	42.8	59.9	73.1	44	45.8
CD49F	0	0	0	0	0	0	0	0	0	0
CD59	83.6	80.1	99.5	87.2	40	42.3	100	100	95.6	95.1
CD90	0.7	4.1	0	0	0	0	88.9	99.9	1.13	5.23
CD117	0	0	0	0	0.3	0.87	0.4	0.59	0	0
CD133	0	0	0	0	0	0	0.3	2.2	0.16	0.21
CD150	0	0	0	0	0	0	0	0	0	0
CD184	99.7	99.9	99.2	100	100	100	99.8	97.1	93.2	91.4
ABCG2	1.13	23.9	1.09	18.1	0.82	9.7	1.17	23.3	1.35	14.1

Table 4.5. Cell surface phenotype of MP and SP cells from 5 PDXs

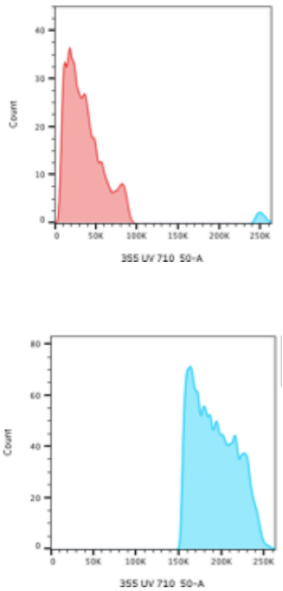
4.2.6 SP cells are much more likely than MP cells to give rise to tumours in NSG mice.

SP and MP cells were isolated from each PDX and collected using the Aria Fusion Cell sorter as detailed in Figure 4.10A. Whilst a small proportion of MP cells were sometimes present in the SP fraction, importantly no SP cells were present in the MP sorted fraction (Figure 4.10B). Tumour formation from isolated SP and MP fractions sorted from all 5 PDXs was monitored and all but one of the 15 SP cell injections resulted in tumour formation while only three out of 15 MP injections led to successful engraftment (Figure 4.10C).

A.



B.



C.

Sample	MP	SP
PDX1	0/3	3/3
PDX2	1/3	3/3
PDX3	1/3	3/3
PDX4	1/3	3/3
PDX5	0/3	2/3

Figure 4.10. SP cells produce tumours in NSG mice.

(A.) Following incubation of PDX cells with Hoechst dye, SP and MP cells were sorted by FACS into separate tubes before re-analysis on a Fortessa to determine the sample purity (B.). (C.) The number of NSG mice that developed tumours following injection of SP or MP cells.

4.2.7 Tumours formed from SP or MP cells *in vivo* give rise to tumours with a minority SP fraction recapitulating the dynamics of the original PDX

An important characteristic of a true CSC population is that they can give rise to heterogeneous populations, composed of both CSCs and differentiated progeny[202]. Tumours produced from SP injected cells were not composed of SP cells alone but also MP cells (Figure 4.11A). Furthermore, the SP fraction in the newly established tumours reached an equilibrium of less than 2% of the bulk tumour despite having been injected as 100% SP cells. In the three mice that produced a tumour from injected MP cells, the tumour that developed contained a SP, although smaller than the SP fraction detected in the original PDX or SP-derived tumour (Figure 4.11 and Table 4.6).

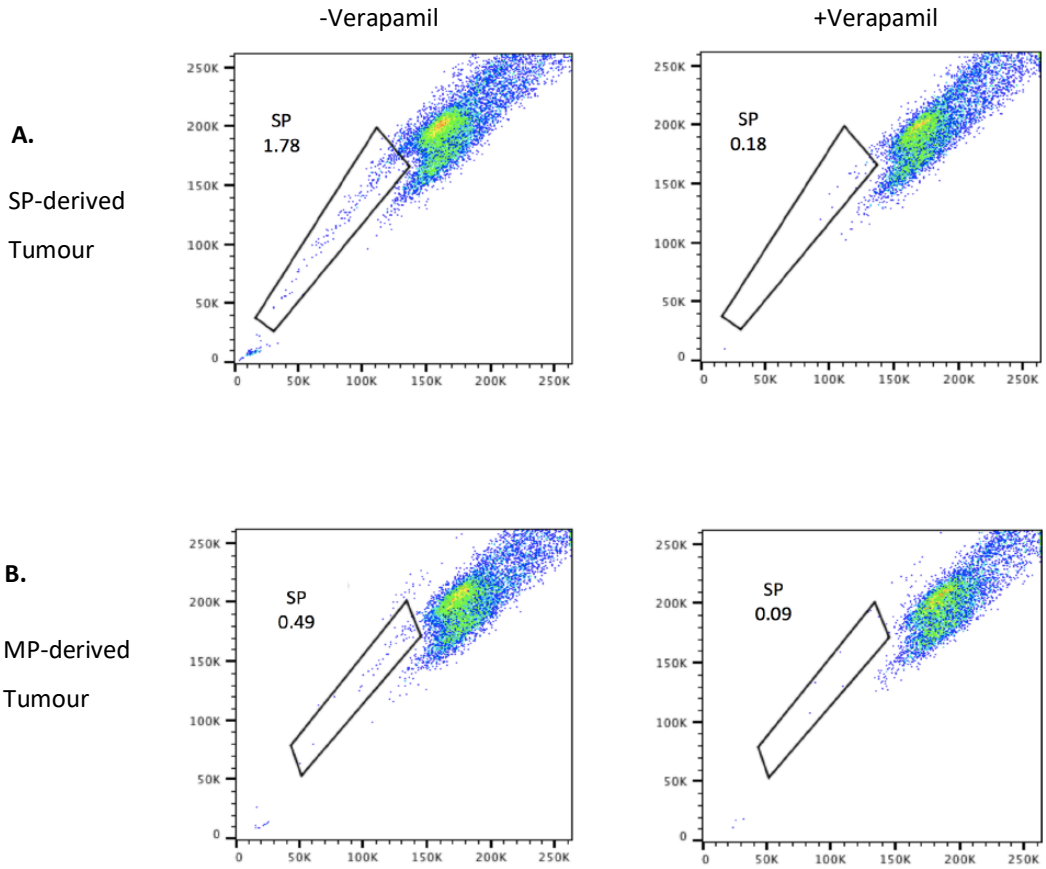


Figure 4.11. Tumours formed from SP and MP injected cells give rise to tumours with a minority SP fraction

Once tumours from SP and MP cells had reached maximum size, cells were incubated with Hoechst dye and analysed for the SP phenotype. The SP fraction in a SP (A) or MP (B)-derived tumour are shown for PDX2.

PDX	% SP in SP-derived tumour	% SP in MP-derived tumour
2	1.78	0.49
3	1.23	0.22
4	2.78	0.65

Table 4.6. Overview of SP fractions in tumours formed from MP cells.

4.2.8 SP cells have increased expression of stemness related genes

If the SP constitutes a CSC population then a difference in gene expression profile compared to MP cells may be expected. Indeed, in the majority of studies where a SP fraction has been isolated, increased expression of genes associated with stemness related pathways have been reported[275,430–432]. In this study, SP and MP cells isolated after passage 1 from each of the five PDXs were analysed by RT-qPCR for expression levels of genes associated with stem cell activity as well as genes involved in B cell development and survival pathways. A number of genes were shown to be significantly increased in SP cells compared to MP cells (Figure 4.12). Of note, there was significant heterogeneity across the five PDXs. Only ABCG2 and ABCB1 were upregulated in the SP of all five PDXs (Figure 4.12). Most genes were only upregulated in some of the PDXs but all SP fractions did show an increased stem-like gene expression profile.

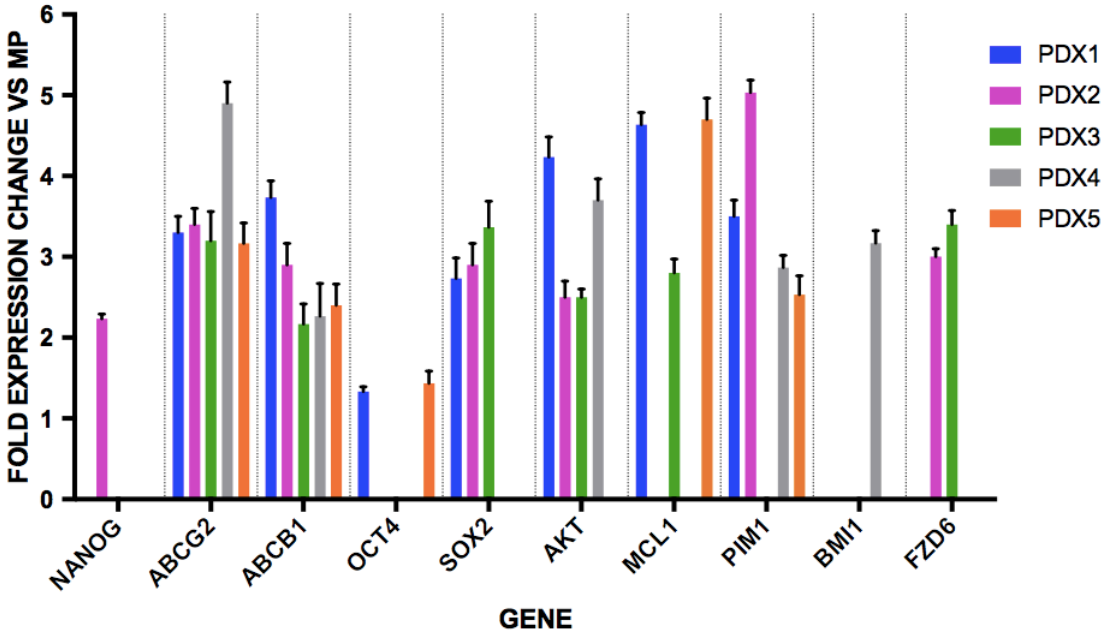


Figure 4.12. Transcript levels of genes associated with stemness and/or survival are increased in the SP compared to the MP fraction of PDX-derived tumours.

After passage, RNA was isolated from SP and MP sorted cells and RT-qPCR carried out for a range of genes. GAPDH was used as a reference control and the fold change in expression levels between SP and MP was determined by calculating $2^{-\Delta\Delta Ct}$*. Data represent the means and standard deviations of technical triplicates. Representative of 2 biological replicates.*

4.3 Discussion

Intratumor heterogeneity caused by genetic, phenotypic or functional differences between cancer cell sub-populations is a considerable clinical challenge and may explain why many cancer therapies fail. Understanding sub-population dynamics is therefore central for both optimization of existing therapy and for the development of new treatment options. Different models have been proposed to explain this heterogeneity including the stochastic model and the cancer stem cell model and in this chapter a number of different methods that have been traditionally used to identify CSCs were explored. PDXs provide an excellent model to carry out these studies as they were shown in chapter 3 to maintain the intrinsic heterogeneity within each tumour.

The initial limiting dilution assays showed that only a small proportion of cells within PDX1-3 could propagate tumours *in vivo*. In PDX2 this was as low as less than 1 in 1000 cells. This clearly highlights that not all cells within the population are capable of propagating the tumour. These cells represent TPCs but further analysis is necessary to determine if they meet all the requirements of a CSC. Limiting dilution assays and xenotransplantation in general has several caveats that must be considered when interpreting data. Cells are being injected into an environment that may not be conducive to tumour growth, even if a cell is a true TIC, and hence cells may be selected for on the basis of their ability to adapt to an environment rather than their TIC potential. The use of the most immunodeficient mice available is an important factor in xenotransplantation and Quintana *et al.* highlighted that the NOD/SCID model underestimated the CSC frequency[304]. They showed that when more immune-deficient NSG mice were used for xenotransplantation in melanoma, the frequency of CSCs was estimated at 25% compared to the 0.1-0.0001% in the NOD/SCID model. NSG mice were used in our study, although the potential for some residual NK cell function as well as defective macrophage and dendritic cell function could impact on engraftment. Furthermore, we utilised matrigel and selected the sub-cutaneous compartment for injection so we may be selecting populations that are conducive to these conditions. Therefore, despite the power of xenotransplantation assays, results should always be considered in the context of these caveats.

The idea of varying levels of tumourigenic potential within the tumour cells was then examined based on cell surface protein expression. Data in chapter 3 outlined the significant level of heterogeneity with respect to cell surface protein expression both within and between the PDX samples. Initial experiments on cell lines indicated that the varying surface protein expression was not just random and rather had significant growth implications. This was translated to the PDXs. The heterogeneity was identified at a phenotypic and functional level rather than genetic but did translate to differential *in vivo* tumourigenic potential of distinct sub-populations. No surface expressed protein alone could identify a population with unique tumour propagating potential although the results seen for $CD9^+$ versus $CD9^-$ populations for both the Ramos cell line and for four PDXs show that $CD9^-$ cells constitute a compartment with significantly increased tumourigenic potential in BL. This agrees with previous studies which have shown that expression of CD9 inhibits the tumourigenicity of human colon carcinoma cells[433] and is inversely associated with cancer metastasis[434] yet contrasts with a study in B-ALL which suggests that $CD9^+$ cells may be candidates for CSCs[435]. The $CD9^+$ and CD9- populations both have tumour propagating potential hence it is unlikely CD9 negative cells constitute a CSC, although their contrasting growth properties compared to the $CD9^+$ population provide evidence of distinct clones within tumours with variable functional properties. This idea of clones with varying levels of fitness was further validated in the other PDX models. For example, the $CD24^+$ population was significantly more tumourigenic in PDX2 as was the $CD44^+$ population in PDX1. CD24 has been identified as a candidate for a CSC marker in human nasopharyngeal carcinoma[436] and in ovarian cancer a population enriched for ovarian CSCs was defined by a $CD24^+$ phenotype[323]. In contrast, the $CD44^+CD24^-$ cell phenotype has been widely reported to define the CSC population in breast cancer highlighting how identifiers of CSCs can vary across cancer types, increasing the complexity of this study area[198,437]. Other cancers that have identified CD44 as a CSC marker include gastric[322,438], colon[439], lung[440] and adult T-cell leukaemia/lymphoma[441]. The $CD49D^+$ fraction which made up the minority of PDX3 had increased tumourigenic potential in mice. CD49D (integrin alpha 4) plays a critical role in leucocyte trafficking, activation, and survival, and also facilitates interactions between leucocytes and stromal cells found in the marrow or germinal centre[442]. In human breast

sarcoma, CD49D^{+} identifies TICs with greater resistance to doxorubicin and in CLL its expression is a negative prognostic indicator for overall survival[442,443]. These are just an example of some of the proteins that showed heterogeneity within and between the PDXs, resulting in varying levels of fitness in tumours and highlight that the heterogeneity within pB-NHL is significant and potentially clinically relevant.

The identification of primitive and stemness-related cell surface proteins on tumour cells in all five PDXs is a particularly interesting observation for BL, a cancer considered to derive from mature germinal centre B cells. CD34 and ABCG2 were expressed on around 1% of cells in all 5 PDXs. CD90 was also expressed on a small fraction of cells in different PDXs except for PDX3 which was almost completely CD90 positive, which is particularly striking. CD34 can be expressed on more differentiated cells such as pro-B cells but CD90 is usually only expressed on CD34^{+}CD38^{-} haematopoietic stem cells and multipotent progenitor cells[429]. Its expression on cells that also express the characteristic BL proteins such as CD20 as well as CD27, a memory B cell marker is of particular note[444]. CD90 is a 25–37 kDa glycosylphosphatidylinositol (GPI)-anchored cell surface protein. Apart from being expressed in haematopoietic precursor cells it is also present on fibroblasts, ovarian cancer cells, endothelial cells and neurons and is involved in MAPK and PI3K siganlling[445]. In gastric cancer and differentiated AML subtypes, CD90 was proposed as a marker to identify CSCs as CD90 expression was high in cells that expressed self-renewal or drug-resistance genes[446,447]. Very little data exist regarding CD90 in BL apart from one study by Ishiura *et al.* which identified strong expression of CD90 in the Akata BL cell line and determined that treatment with anti-Thy1 antibody inhibits their proliferation more strongly than Rituximab[448]. A study of expression of CD90 antigen on pre-B cell ALL with t(9;22) indicated that it correlated with the proliferative status of the leukaemia cells, not their phenotypic immaturity[449]. The mechanisms and significance of CD90 expression in mature pB-NHL remains to be elucidated although its expression, and the increased fitness of CD90^{+} cells versus CD90^{-} cells in PDX4, again highlight an intriguing level of heterogeneity both within and between patient samples. The expression of ABCG2 in the PDXs is another significant result as this protein has been associated with CSCs, malignant behaviour and drug resistance in a broad array of cancers[147,450–452]. ABCG2 is expressed in several normal

tissues but its expression is sharply down-regulated during hematopoietic stem cell differentiation and is expressed at a very low level in mature cells compared with progenitor cells[333]. ABCG2 expression is detected in several cancers and generally this expression is associated with negative prognosis and multi-drug resistance in tumours, for example lung, oesophageal and breast cancers[333,453,454]. A study by Kim *et al.* evaluated ABCG2 expression in 67 cases of DLBCL. High levels of ABCG2 on tumours correlated with shorter OS as well as Sonic Hedgehog protein levels – a key stem cell associated signalling pathway[455].

The identification of progenitor and stem related markers in the PDXs leads to numerous hypotheses and requires further study. The initiation of carcinogenesis in BL could be in a more primitive cell than previously thought and, akin to models of leukaemogenesis, potential pB-NHL stem cells could derive either from HSCs or multipotent progenitors. Even though advances have been made in the understanding of CSCs in leukaemias and solid tumours, this theory has not been studied in mature lymphoid malignancies until recently. Evidence has emerged of apparent transmission of premalignant lymphoid cells from 1 individual to another in at least seven cases including FL, MCL, T-cell lymphoma, mucosa-associated lymphoid tissue lymphoma and CLL[283–288]. Despite these advancements in knowledge, few studies have been conducted in this area for DLBCL and BL and for those that exist, the evidence is not as conclusive[296,297]. Alternatively, the existence of stem-related characteristics could be due to epigenetic changes during tumorigenesis whereby more differentiated cells are reprogrammed to acquire stem cell features; in evidence epigenetic regulation of ABCG2 has been widely reported[456,457]. Regardless as to whether any of these proteins identifies a CSC in pB-NHL, their presence are important observations in the study of intra-tumour heterogeneity.

Due to the level of heterogeneity between our samples with regards to cell surface expression of proteins and the difficulty in pinpointing a CSC population by surface phenotyping alone, a functional assay that has been shown to denote cells with stemness was employed. Following development of the SP assay by Goodell *et al.* in 1996 this technique has been employed in a whole host of studies to identify stem-like cells in healthy and malignant tissue. Despite a lull in the study of SP cells in the early

part of this decade, a study in 2016 by Boesch *et al.* which identified SP cells in 12/13 unrelated human cancers, renewed interest in this area[458]. After careful optimisation of the assay in this study we were able to detect a SP fraction in all five of our PDXs. The first evidence that these may indeed constitute a CSC population was the ability of the SP cells to propagate tumours in NSG at a much higher rate than MP cells. The tumours formed from SP cells also underwent asymmetric division with the presence of both SP and MP cells in the ensuing tumour. Interestingly, the population of SP cells in the newly developed tumour reduced from 100% SP when injected, to approximately 2% after tumour formation. Hence, not only do these cells produce non-SP cells but the tumours also evolve towards a state where SP cells make up the minority of the tumour. However, some injected MP cells (one in three of three different PDXs) produced tumours. If the SP fractions represented a true CSC compartment then the non-SP cells should not have long-term replicative capabilities. The possibility that some SP cells passed into the MP fraction during sorting cannot be disproven although the gates were stringently applied to avoid this occurring and doublets were eliminated in the gating strategy. Alternatively, the MP population may be contaminated with some SP cells due to the very sensitive incubation technique. It is possible that some cells, with the ability to, had not yet effluxed dye and hence were present in the 'MP gate'. Interestingly the 3 tumours formed from MP cells had a small fraction of SP cells present on tumour formation and this population was able to produce a tumour on further passage, highlighting long-term propagating potential. A further interpretation of this result draws on an idea mentioned above whereby the CSC population may be more transient than previously thought. Non-SP cells may be able to acquire stemness characteristics due to epigenetic changes. An unstable CSC phenotype can lead to phenotypic conversions due to cell-intrinsic or microenvironmental cues[415]. Indeed, as discussed in chapter 1, numerous studies have highlighted the ability of the TME to induce a stem cell phenotype in mature B cell malignancies, in particular it has been shown that stromal cells activate MYC, EZH2, OCT4, NANOG and SOX2, key drivers of the stem cell phenotype[234,235]. Some other studies also support the hypothesis of inter-conversion of phenotype if conditions are correct[459]. In particular, a study in DLBCL highlighted that conversion of cell states between SP and non-SP cells was possible and was mediated

by exosome-mediated Wnt signaling[230]. It is therefore possible that in this study, cells within the MP injected population were able to convert to SP cells due to a particular environmental or cell-intrinsic cue that influenced the transition. Mitsutake *et al.* reported a similar finding in their study of thyroid cancer; 6/8 of their non-SP injected mice produced tumours, and although the SP cells were more clonogenic than non-SP cells, the clonogenic ability of non-SP cells was not zero. Their interpretation was that CSCs are enriched in the SP but not exclusively present[460]. In fact similar results have been seen in a whole host of recent studies including those in cervical cancer, laryngeal squamous cell carcinoma, mantle cell lymphoma and hepatocellular carcinoma[432,461–463]. In all these studies, some tumours formed from non-SP cells, although they were reported to be significantly less frequent or tumourigenic than their SP counterparts.

Kreso and Dick consider the prospective identification and isolation of the cells as a requirement for a population to be considered a CSC[202]. After identifying a SP fraction in each of the PDXs studied here and determining their tumour initiating capabilities, a more specific cell surface phenotype was investigated to enable identification of the SP fraction. As ABCG2 expression is a defining feature of the SP fraction we initially looked at co-expression of this surface protein with other stem cell associated cell surface proteins in PDX4. These data highlighted that $ABCG2^{+}$ cells were all CD24, CD90 and CD49D positive. Identifying distinct expression profiles associated with the $ABCG2^{+}$ cells in PDX4 provided evidence that it may be possible to determine a specific phenotype for the CSC-like cells in BL. However, when phenotyping of the separated SP and MP fractions was conducted for all five PDXs, although ABCG2 expression was significantly higher in the SP cells, it was not exclusive to the SP compartment. In addition, markers determined as being specific to the ABCG2 compartment in PDX4, such as CD133, were also found in both MP and SP fractions. Despite this, there were significant increases or decreases in the proportion of cells expressing certain proteins in the SP of all five PDXs. CD9 expression is on a lower percentage of cells in the SP of all PDXs that express it which is particularly interesting considering the increased tumourigenic potential of $CD9^{-}$ PDX cells *in vivo.* Studies have shown that there is an inverse correlation between CD9 expression level and tumour metastasis in solid tumours[434] and Yoon *et al.* also showed that down-regulation of CD9 expression correlates with tumour progression in

B cell lymphomas[464]. CD90 which is expressed on tumour cells of PDXs 1,4 and 5 is expressed on a higher proportion of SP cells, indeed all cells in the PDX4 SP are $CD90^+$. CD49D expression was expressed on a higher fraction of the SP of three out of five PDXs which is particularly interesting considering that $CD49D^+$ cells from Ramos and PDX3 had significantly increased tumourigenic potential *in vivo*. The increase in primitive cell marker expression is also an important observation and in particular CD34 is expressed on more cells in all SP fractions. The fact that ABCG2 and other primitive cell markers are increased in SP cells suggests that the SP phenotype is enriching for CSCs, but may not encompass the entire CSC population – the previous discussed growth of three tumours from MP cells may support this theory. Alternatively, the CSC population may indeed lie specifically within the SP but be defined by a certain cell surface phenotype involving a combination of markers. Nonetheless, the fact that cell surface expression phenotypes determined earlier in our studies to confer increased fitness *in vivo* are associated with the SP in the PDXs, is an encouraging indicator that we are successfully prising apart the heterogeneity within these PDXs and most importantly identifying populations with functional importance.

To determine further characteristic features of the SP fractions that may be contributing to their increased tumourigenic potential, RT-qPCR expression analysis of a host of genes was carried out. Genes chosen included ones associated with stemness properties, as well as anti-apoptotic and therapy resistance genes which have been shown to be upregulated in SP cells in other studies[230,459,465]. As expected, ABCG2 and ABCB1 were both expressed by a higher percentage of SP cells for all 5 PDXs. The primitive cell-associated gene NANOG was only upregulated in the SP of PDX1 while OCT4 was only expressed in two PDXs and its increased expression in the SP was minor. SOX2 was expressed by all PDXs and was upregulated in the SP fraction for three of them. Apart from ABCG2 and ABCB1, no single gene was differentially expressed in all 5 PDX SP fractions compared to the MP although the anti-apoptotic proteins MCL1 and PIM1 were upregulated in the SP of three and four different PDXs, respectively. Both of these have been shown to be important in B-NHL with PIM1 known to cooperate with BCL6 to promote the development of lymphoma[466] while targeting of MCL1 kills MYC-driven mouse and human lymphomas even when they bear mutations in *TP53*[467].

Furthermore, targeted genomic sequencing of paediatric BL has identified recurrent alterations in the MCL1 gene and other anti-apoptotic genes[468]. FZD6 was upregulated in the SP of two PDXs; this gene is a member of the 'frizzled' gene family, which encode receptors for Wnt signalling proteins. In liver cancer, SMARCA4, a SWI/SNF family member mediates transcriptional activation of FZD6 initiating self-renewal of TICs and SMARCA4 is commonly mutated in BL[52,469].

Finally, expression levels of AKT, a member of the PI3K pathway, were increased in the SP fractions of 4 PDXs. This is an interesting insight as the PI3K pathway has been reported to have a role in regulating ABCG2 expression[276]. The fact that there is heterogeneity with respect to the expression profile of the SP fraction in the five PDXs is again an interesting observation – although they all have signatures associated with upregulation of stemness and anti-apoptotic programs at the transcript level, there is no overlapping signature in all five, at least with respect to this gene set. In a recent study by Wagener *et al.*, a PluriTest algorithm was used to investigate traces of pluripotency in germinal centre derived B cell lymphoma and determined that BL cells more closely resemble human pluripotent stem cells than non-molecular BLs and intermediate lymphomas do, arguing for a greater role of stemness in BL compared to other mature B-NHLs[470]. Given that the main characteristic that distinguishes BL from these other entities is deregulated MYC expression it suggests a possible role for this gene in promoting stemness. In the context of our data this is plausible as the expression analysis above highlighted that cell cycle progression and proliferative activity are important signatures of the SP.

The observed heterogeneity in expression levels across the PDXs may allude to the fact that the identification of a distinct CSC with the same features in each patient may not be possible. As outlined by Kreso and Dick in their unified model of cancer - genetics, epigenetics and the microenvironment can each influence stemness and hence may not converge on the same characteristics in all tumours - which will in turn influence clinical outcome[202]. Indeed, in a study of the TICs in 16 AML samples, it was shown that the cell fraction(s) that contained TICs, identified by xenotransplantation assays, were variable for each sample[319]. However, the expression profile for all samples converged on a

stemness signature with the TIC signature prognostic for overall survival across a wide spectrum of AML patients. The idea that the CSC compartment might not be the same across patients is another daunting prospect when it comes to translating the CSC model to patient therapies. However, by adapting a personalised medicine approach, with the use of PDXs, the drivers of stemness in individual tumours can be elucidated and targeted. For example, BMI1, which was upregulated in one of our PDX SP fractions, has been strongly linked to self-renewal and implicated in the maintenance of stem cells in several tissues[471–473]. It was found that human colorectal TIC function is dependent on BMI1 and down-regulation of its expression inhibited the ability of colorectal TICs to self-renew, resulting in abrogation of their tumorigenic potential[474]. Treatment of primary colorectal cancer xenografts with a small molecule BMI1 inhibitor resulted in colorectal TIC reduction with long-term and irreversible impairment of tumour growth. Hence, utilising PDXs to determine the CSC or other populations driving relapse or resistance in pB-NHL and devising targeted treatment options may be a rational approach in these difficult to treat cases.

Chapter 5. Investigating SP growth dynamics in vitro and potential SP-targeting therapeutics

5.1 Introduction

The CSC hypothesis states that within a given tumour cell population a small subset of cells are capable of propagating the tumour and importantly can give rise to both tumourigenic and non-tumourigenic cells[200]. Hence, by definition, if the SP compartment denotes the CSC then these characteristics should be met, whether this SP population is in a primary tumour or a cell line. Whilst a SP has been identified in all 5 PDXs, the inability to grow these tumour cells *ex vivo* meant further analysis of the growth dynamics of these populations was difficult. Therefore, the presence of a SP within paediatric BL cell lines was investigated to determine if they meet the characteristics of a CSC. Most of the literature on the SP assay has been conducted using cell lines, with a recent study showing the presence of SP fractions in cell lines representing 12 different cancer types[458]. In addition, many studies have highlighted the increased tumourigenic potential of cell line-derived SP cells *in vivo* compared to SP-depleted cell lines[475,476,477,478]. Moreover, *in vitro* studies have demonstrated that SP cells have increased viability compared to MP in cell lines derived from hepatocellular carcinoma, squamous cell carcinoma, cervical cancer, osteosarcoma and Anaplastic Large Cell Lymphoma[430,431,461,462,465,478].

In the previous chapter, the differential expression of potentially targetable proteins associated with stemness, anti-apoptotic pathways and PI3K signalling in the SP fraction of PDXs was determined. In the literature, a host of signalling pathways have been shown to be involved in maintenance of the CSC phenotype. For example, the Notch signalling cascade is a major pathway involved in numerous critical cellular processes including stem cell maintenance, progenitor cell proliferation and differentiation[255]. Furthermore, activation of Notch4 and Notch1 has been noted in breast CSCs, and inhibition of these receptors, particularly Notch4, leads to significantly decreased mammosphere formation *in vitro* and complete inhibition of tumour initiation *in vivo*[479]. Wnt signalling is another highly deregulated pathway in cancer and it has been shown to be involved in self-renewal and maintenance of stem cells and CSCs in several tissues

including skin, intestine and mammary gland[480]. A minor subpopulation of breast TICs associated with drug resistance was identified in human breast cancer; *in vitro* and *in vivo* studies showed that suppression of Wnt signalling inhibited sphere- and colony-formation of primary breast tumour cells and blocked tumour growth in murine models[481]. Several studies also suggest the involvement of the PI3K/AKT/mTOR pathway in maintenance of CSC features. For example, in breast cancer, inhibition of PI3K or AKT activity reduced generation and growth of CD44/CD24 mammospheres, leading to stem cell/mesenchymal phenotype loss and recovery of epithelial-like markers[482]. In prostate cancer, PI3K/AKT/mTOR pathway deregulation is associated with CSC maintenance; moreover, prostate CSCs are resistant to selective mTOR inhibitors[483]. As such, there are potentially many therapeutically-actionable targets for the elimination of populations with stemness in cancer. Given the inability to grow and assay primary patient tumours *in vitro*, SP fractions were isolated from paediatric BL cell lines in order to conduct *in vitro* studies of the growth dynamics of these cells and their response to drugs.

Aims of this chapter

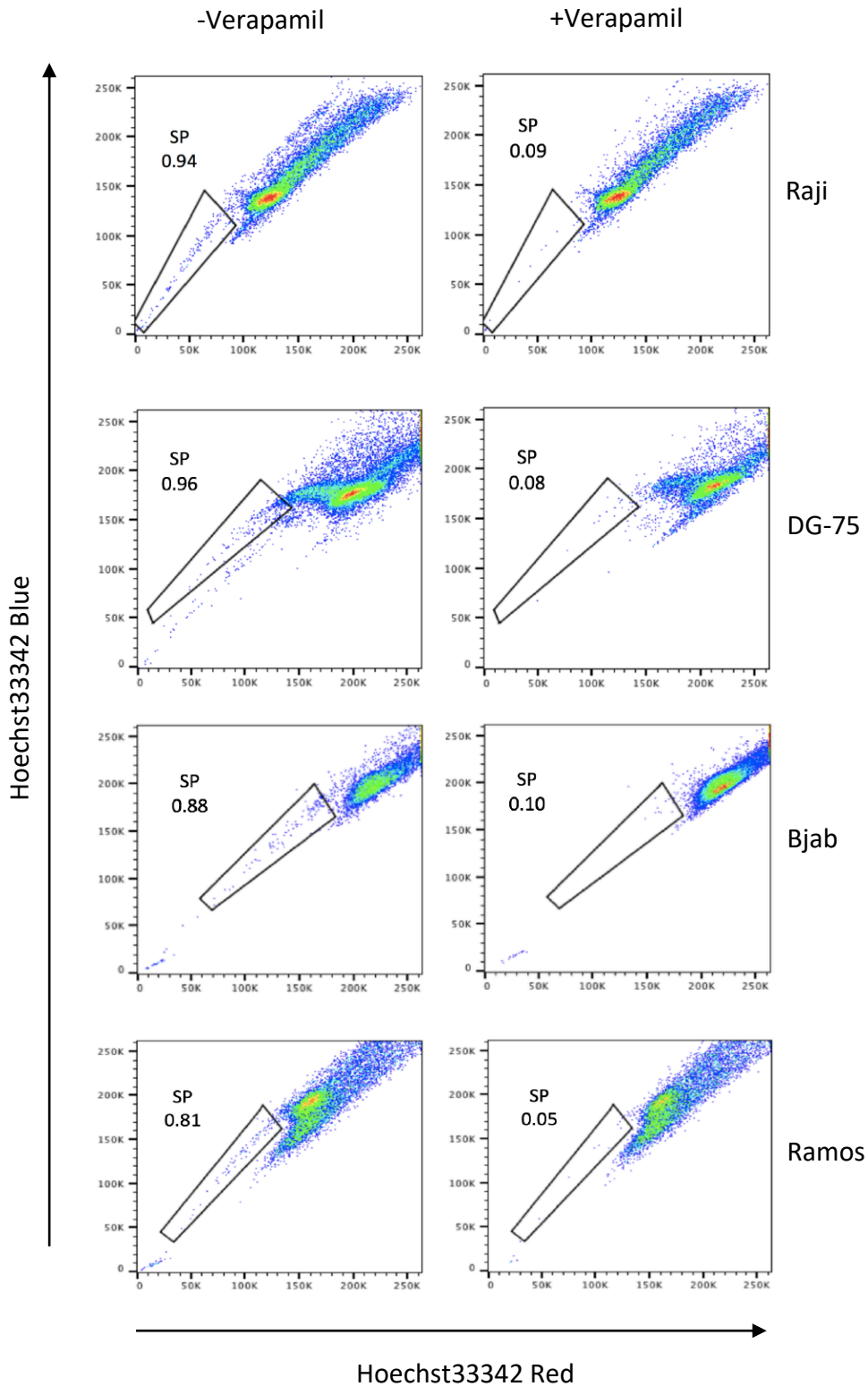
- Determine if a SP is present in paediatric BL cell lines and optimise the procedure for its identification
- Determine if cell line SP cells have CSC properties
- Compare the growth potential of SP and MP cells *in vitro*
- Compare the sensitivity of SP and MP cells to conventional chemotherapies
- Determine if the survival of SP cells can be inhibited by targeted drugs.

5.2 Results

5.2.1 A Side Population exists in BL cell lines

Optimisation was carried out in the same way as described for the PDXs as detailed in chapter 4. Briefly, DG-75, Ramos, Raji and BJAB cell lines were incubated with increasing concentrations of Hoechst dye (2-10ug/ml) for varying time intervals (60, 90, 120 mins). Initial optimisation was conducted with cells cultured in 10% FBS revealing very small SP fractions of 0.2% or less (data not shown). Several studies have identified the importance of different culture conditions to the SP phenotype and the possibility for variation in the population depending on confluency, serum levels and hypoxia[484]. When the FBS concentration was altered to 2% for seven days culture before analysis, the SP fraction was much more abundant, a feature that has been noted in other SP optimisation studies[356]. As these conditions may better represent the tumour microenvironment *in vivo* where it is known that nutrient starvation, hypoxia and poor blood supply are common, 2% FBS containing media was used in subsequent SP experiments[485,486]. To maintain consistency between experiments, other variables such as cell density and assay conditions were kept constant. Each of the four cell lines had an SP of between 0.8-1% (Figure 5.1A) although there was slight variation in the exact percentage with each assay. The average of five separate experiments per cell line identified Raji to have the greatest variation in SP percentage with an average SP value of 0.95% and a standard deviation of 0.08% (Figure 5.1B).

A.


-Verapamil
+Verapamil
SP
0.94
SP
0.09
Raji
SP
0.96
SP
0.08
DG-75
SP
0.88
SP
0.10
Bjab
SP
0.81
SP
0.05
Ramos
Hoechst33342 Blue
Hoechst33342 Red

B.

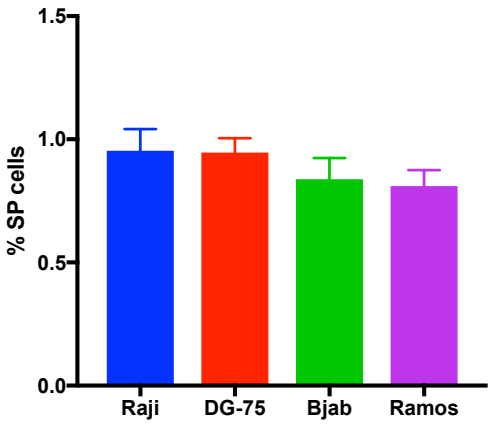


Figure 5.1. BL cell lines contain a SP

(A.) Data representative of the SP detected in each of the cell lines compared to Verapamil-treated control cells. (B.) The SP ranges in quantity for each cell line in independent assays. Data represent the means and standard deviations of five independent experimental replicates conducted for each of the indicated cell lines.

5.2.2 SP cells have greater proliferative capabilities than MP cells *in vitro* and support the growth of the MP; SP-derived soluble factors drive MP growth

In order to investigate the proliferative properties of SP and MP cells in vitro, 10,000 of each cell type were sorted and cultured over a 10 day period. A gating strategy for sorting of the cell populations was developed in order to minimise contamination within cell populations (Figure 5.2). Unstained, whole cell populations were also passed through the cell sorter to ensure they experienced the same conditions of stress. Importantly, cell viability analysis was carried out on the SP, MP and unsorted population, which had also passed through the sorter, and there was no significant difference in cell viability in the 3 populations. This suggests that the staining procedure with Hoechst dye does not have a negative effect on cell viability (Figure 5.2 B) The live cell count was determined every second day for each population in triplicate samples following trypan blue exclusion (Figure 5.3). In general, the isolated SP and whole cell populations begin to proliferate immediately, whereas MP cells stagnate. For example, over the 10 day period the Bjab whole cell population and SP underwent almost 4 population doublings. This is slightly less proliferative than normal for this cell line which

usually doubles every 50 hours (i.e. over 4.5 population doublings in 10 days). These data suggest that the sorting process may have a slightly negative impact on cell growth due to cellular stress. In comparison, the average MP cell number was 19300, or almost one population doubling in the same timeframe, showing that the population has significantly limited proliferative capabilities (Figure 5.3). Similar results were seen for the Raji and DG-75 cell lines. However, the Ramos cell line, whilst following a similar trend for the whole cell population and SP, did show a higher proliferative capacity for the MP cells but still significantly less than the Ramos unsorted whole cell population or isolated SP (p<0.005, Figure 5.3).

Considering that the growth kinetics of the whole population of cells was almost as expected but that the MP cells had low proliferative abilities when isolated, we hypothesised that the MP cells required the SP for effective proliferation. This premise was investigated by re-combining SP and MP cells in the same well of a tissue culture plate. SP cells (1000) were added to 9000 MP cells and the total live cell count monitored by trypan blue exclusion. Although this is a higher ratio of SP cells than in the normal cell line, 1000 cells were chosen to minimise the negative impact of low confluency on SP cell viability. This culture set-up resulted in a growth rate similar to that of the normal unsorted cell line, i.e. the growth potential of the MP cells had recovered in the presence of the SP cells (Figure 5.3). To further determine if this recovery was due to direct cell contact or soluble factors secreted by the cell populations, we utilised transwell plates whereby two cell populations are separated by a permeable membrane that allows free movement of soluble factors but not cells between compartments. As was noted on re-combining SP and MP cell fractions, free flow of media and soluble factors between the cell populations supported growth of the MP cells (Figure 5.3). Hence, it is likely that soluble factors produced by the SP are responsible for driving MP cell growth.

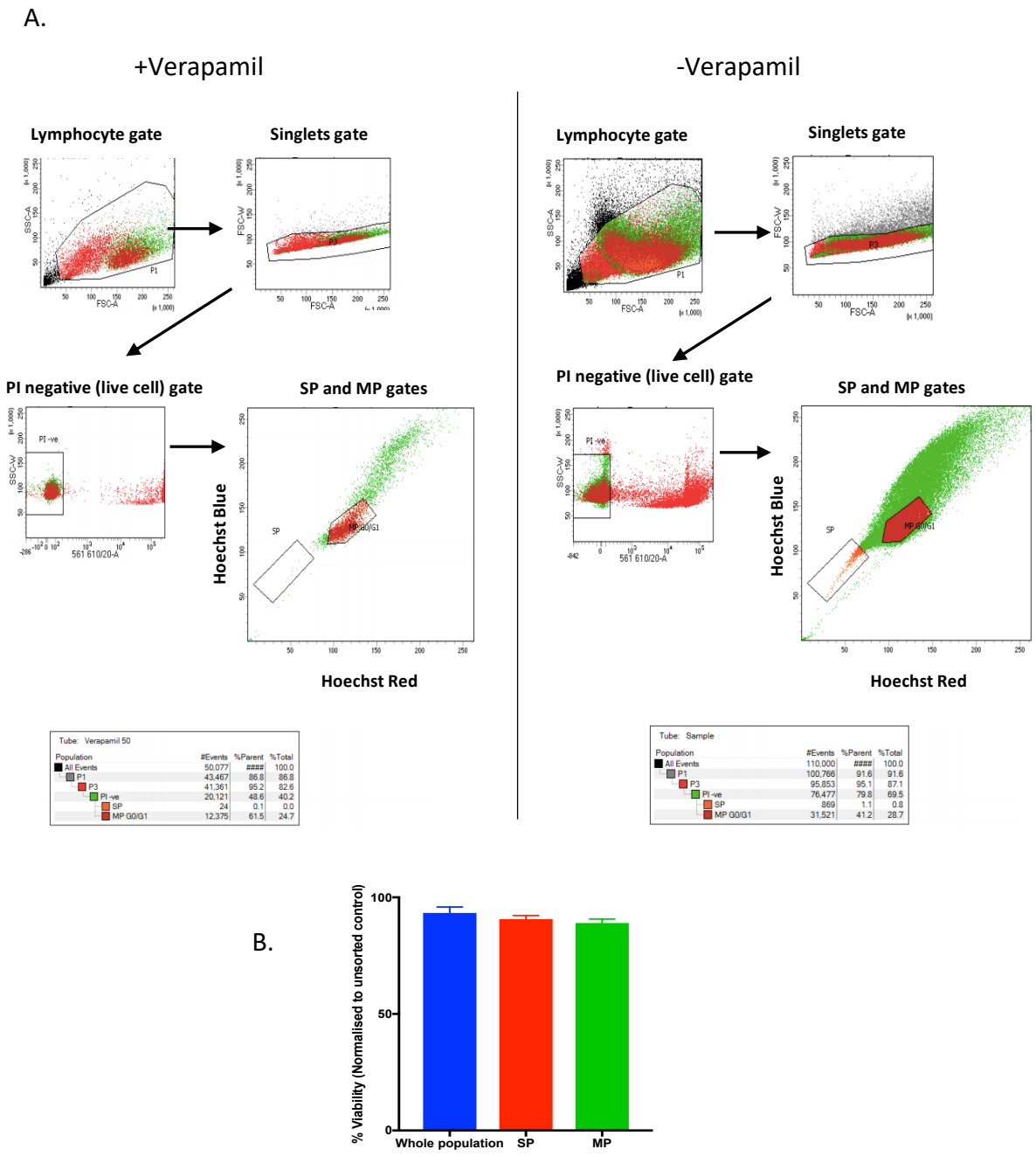


Figure 5.2. The gating strategy for sorting cell line-derived SP and MP cells by FACS

(A.) *Following processing in either the presence or absence of verapamil, the main lymphocyte population was gated, followed by the singlet population and then the PI negative population before analysis for SP cells. The SP gate was devised based on the position of the Verapamil treated cells (left panel). (B.) CellTiter-Blue viability assay was performed on whole populations cells, as well as SP and MP cells from the Raji cell line which had been incubated with Hoechst dye. All populations passed through the sorter. Viability was normalised to unsorted control. Means and standard deviations of technical triplicates shown. Representative of 3 replicate experiments.*

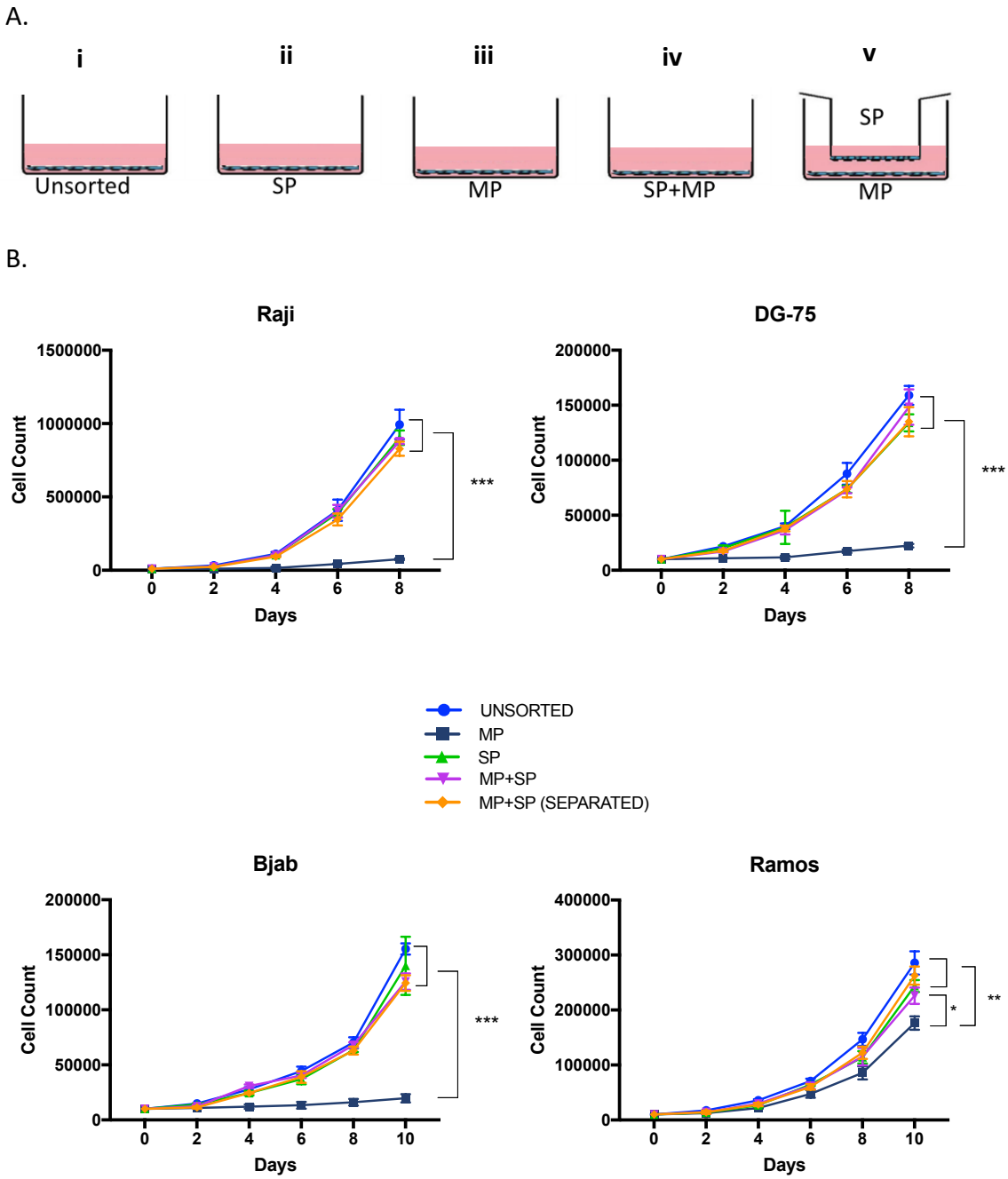


Figure 5.3. The SP has significantly increased proliferative capabilities in vitro and supports the growth of MP via production of soluble factors

*(A) MP and SP cells were separated by FACS before (co-)culture as indicated: 10,000 of each cell type was plated per well for (i-iii); 9000 MP and 1000 SP were plated in (iv); 9000 MP cells were plated in (v) with 1000 SP cells in the well insert. (B) Cell growth was monitored every 2 days following trypan blue exclusion. Data are representative of the means and standard deviations of technical triplicates, ***p<0.001, **p<0.005, *p<0.0. Representative of three biological replicates.*

5.2.3 Exosomes isolated from SP cells do not recover MP growth

In order to try and elucidate the soluble factors that may be driving MP cell growth, exosomes were isolated from SP cells and exogenously applied to the MP population. Exosomes have been recognized as mediators of inter-cellular communication among different cell populations and by transfer of signalling molecules such as proteins, lipids and RNAs between different cell types, they are implicated in both physiological and pathological processes[487]. Several studies in recent years have highlighted that exosomes released by CSCs mediate cancer progression in different model systems[488]. Most notably, exosomes were found to mediate population equilibrium of lymphoma stem like cells and non-stem cells in DLBCL[230]. Following 72 hours in culture, exosomes produced by isolated Bjab and DG-75 SP cells were harvested and their presence verified by transmission electron microscopy (TEM) before supplementing the media of sorted MP cells (Figure 5.4); their growth rate was monitored following trypan blue exclusion over 6 days. The growth rate of MP cells did not increase in the presence of SP exosomes (Figure 5.5), suggesting that they are not the soluble factors responsible for the previously observed MP growth in the presence of SP-conditioned media.

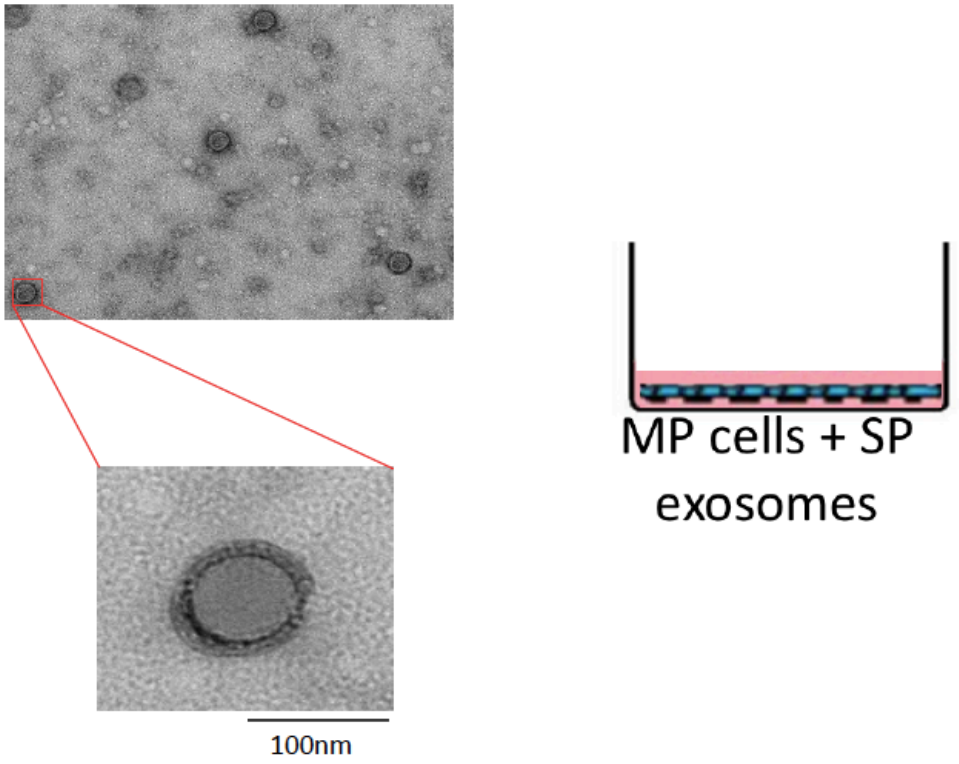


Figure 5.4. Confirmation of exosome presence following isolation from Bjab or DG-75 conditioned media prior to supplementation of MP culture media

Isolation of exosomes by ultracentrifugation was confirmed by Transmission Electron Microscopy before supplementing MP cell media.

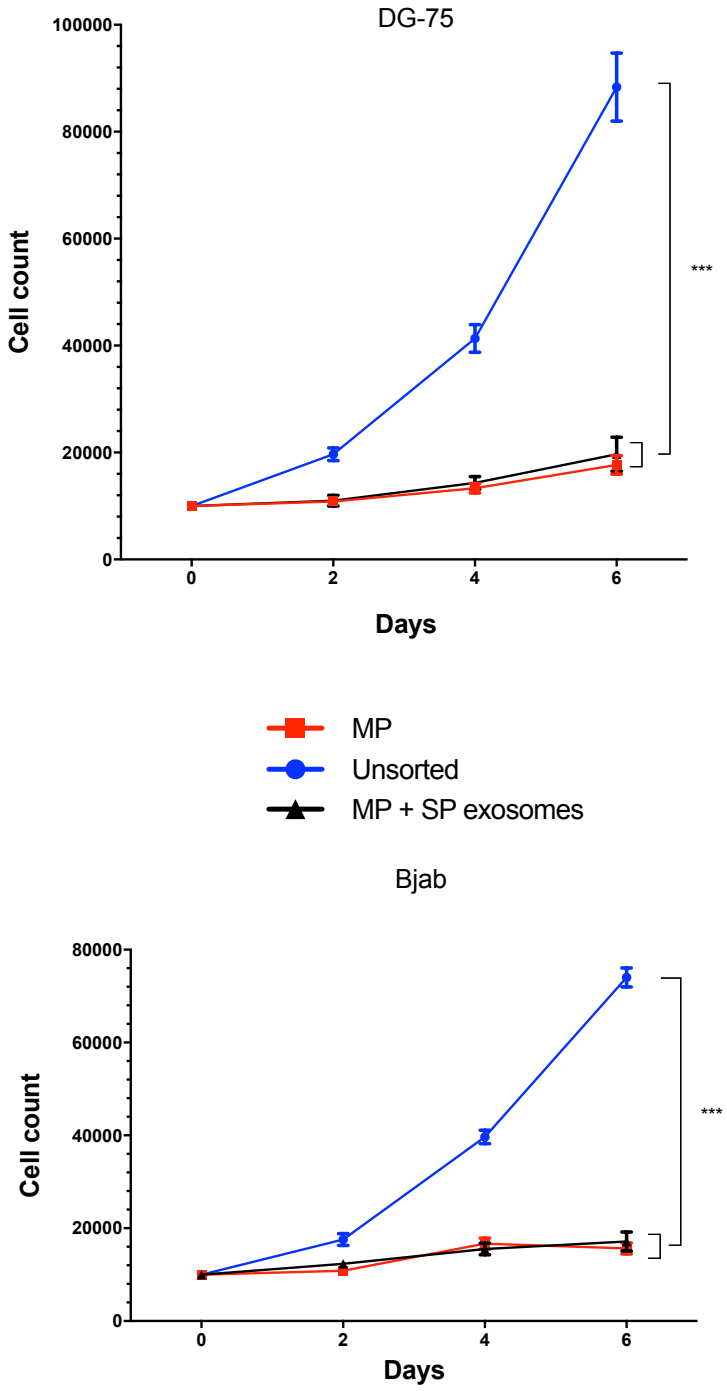


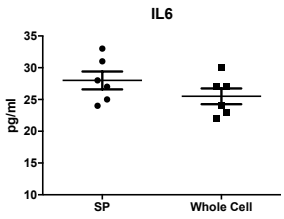
Figure 5.5. SP-exosome enriched media does not recover MP cell growth

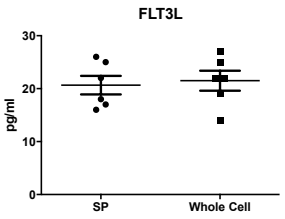
*Exosomes isolated from SP-conditioned media of the indicated cell lines were added to culture media and the growth rate of 10000 MP cells monitored following trypan blue exclusion. Data represent the means and standard deviations of technical triplicates. ***p<0.001*

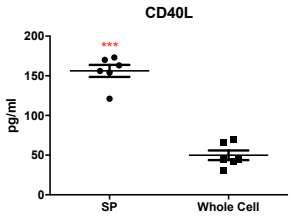
5.2.4 The levels of certain cytokines produced by SP and MP cells differ

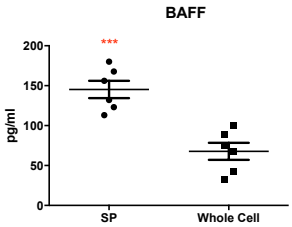
An obvious method of communication between cells is via cytokines and growth factors. Indeed, numerous studies have highlighted the impact of cytokine expression patterns on cancer cell stemness[226–228]. In order to investigate the cytokine profile secreted by the SP, 20,000 SP cells from each cell line were sorted by FACS into wells in triplicate. As the MP does not grow successfully *in vitro* the whole cell population was used for comparison of the cytokine profile given that they consist of ~99% MP cells. After 72 hours in culture, supernatant was separated by centrifugation and cytokine production determined using an Immunoplex assay. This offers broad range and sensitivity for cytokine detection and the panel employed detects cytokines usually produced by haematopoietic stem cells and B lymphocytes. Cytokines that were not detected in the supernatant of any of the cell populations/cell lines are not shown. Of note, there was no cytokine produced at a significantly higher level by either the SP or MP that was common to all of the cell lines. The Bjab SP produces the most differentially expressed cytokines with CD40L, BAFF and IL-10 significantly higher in the SP compared to the whole cell population (Figure 5.6). Interestingly, IL-10 was also significantly increased in Raji SP and BAFF in the DG-75 cell line SP, while IL11 production by DG-75 SP cells was also significantly higher. As such, there may be heterogeneity with respect to the underlying driver of the SP phenotype of the three cell lines, if cytokines are the responsible mechanism. However, an exhaustive list of cytokines was not investigated and so further studies using a wider panel may identify common cytokines that are significant to all three cell lines. Alternatively, a combination of factors that activate common downstream pathways in the MP cells might be responsible.

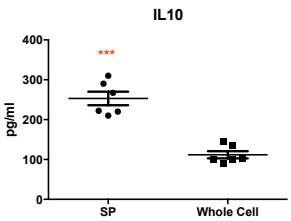
Bjab

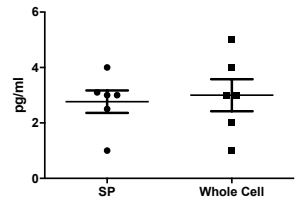

IL6
pg/ml
35
30
25
20
15
10
SP
Whole Cell

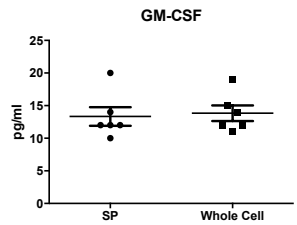

FLT3L
pg/ml
30
20
10
0
SP
Whole Cell

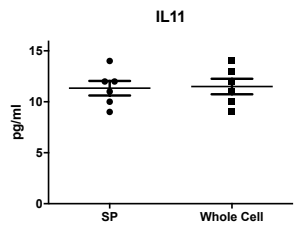

CD40L
***
pg/ml
200
150
100
50
0
SP
Whole Cell

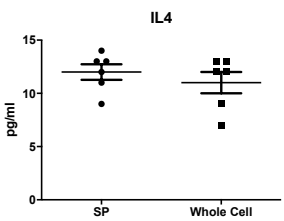

BAFF
***
pg/ml
200
150
100
50
0
SP
Whole Cell


IL10
***
pg/ml
400
300
200
100
0
SP
Whole Cell

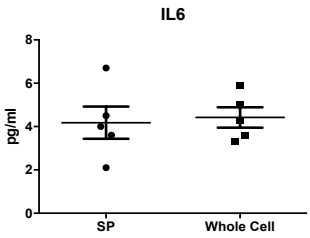

IL34
pg/ml
6
4
2
0
SP
Whole Cell

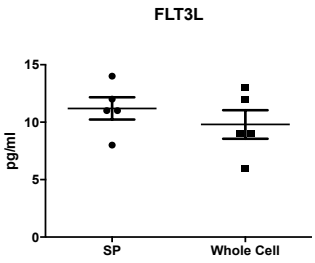

GM-CSF
pg/ml
25
20
15
10
5
0
SP
Whole Cell

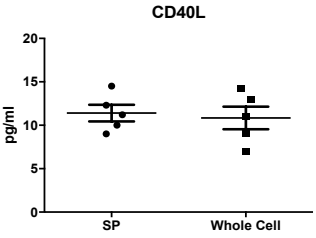

IL11
pg/ml
15
10
5
0
SP
Whole Cell

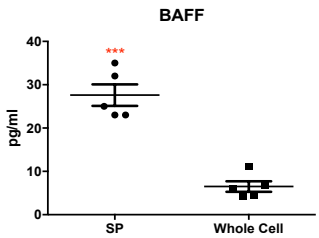

IL4
pg/ml
15
10
5
0
SP
Whole Cell

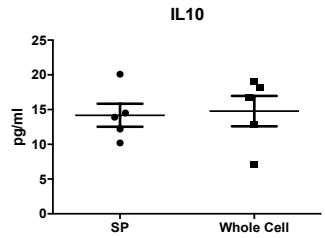
DG-75

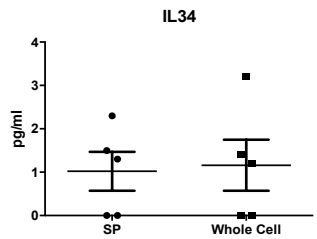

IL6
pg/ml
8
6
4
2
0
SP
Whole Cell

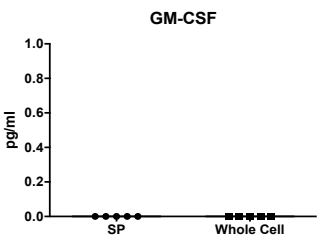

FLT3L
pg/ml
15
10
5
0
SP
Whole Cell

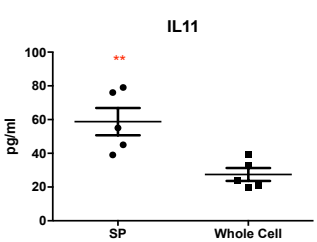

CD40L
pg/ml
20
15
10
5
0
SP
Whole Cell

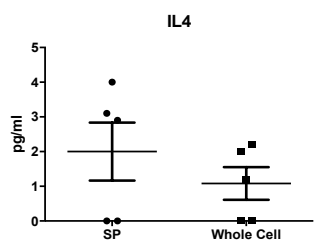

BAFF
***
pg/ml
40
30
20
10
0
SP
Whole Cell


IL10
pg/ml
25
20
15
10
5
0
SP
Whole Cell


IL34
pg/ml
4
3
2
1
0
SP
Whole Cell


GM-CSF
pg/ml
1.0
0.8
0.6
0.4
0.2
0.0
SP
Whole Cell


IL11
**
pg/ml
100
80
60
40
20
0
SP
Whole Cell


IL4
pg/ml
5
4
3
2
1
0
SP
Whole Cell

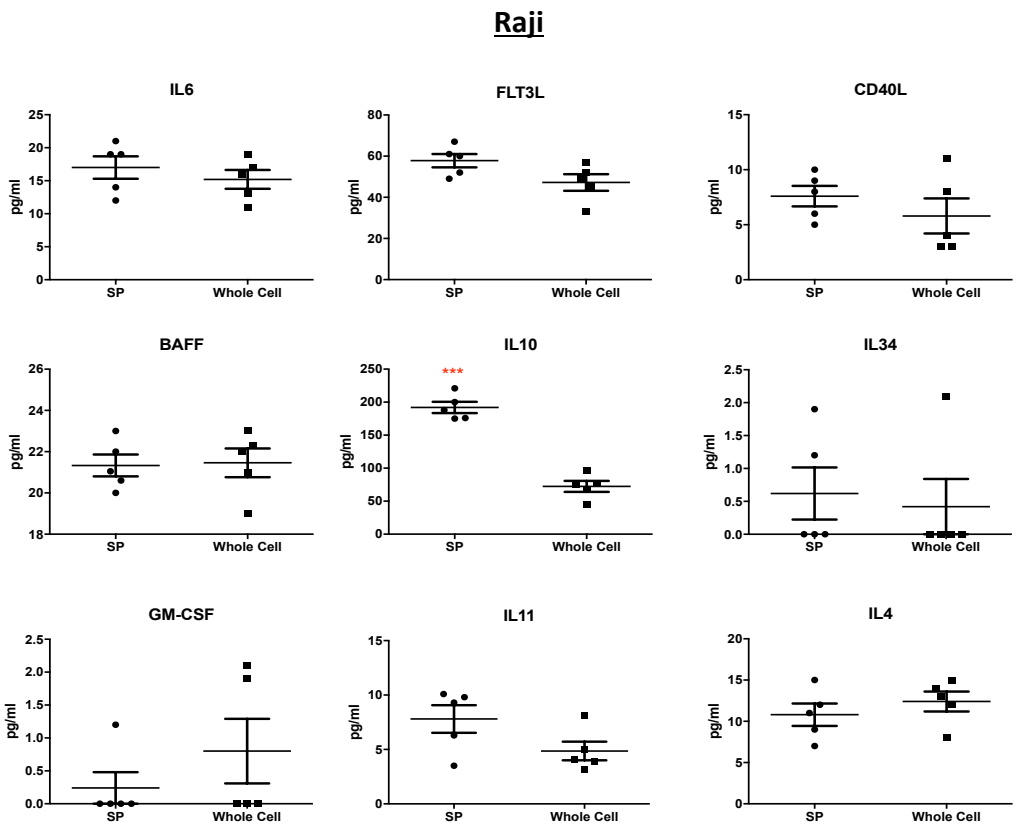


Figure 5.6. Differential expression of cytokines in SP and whole cell populations.

*The concentration of soluble analytes of the respective cytokines present in the supernatant of the SP or whole cell line was determined for Bjab, DG-75 and Raji cell lines by Biolegend LEGENDplex™ immunoassays. Data show individual data points of 5 technical replicates with the mean and standard error of the means for 1 representative experiment of 3. *P<0.05, **p<0.005, ***p<0.001*

5.2.5 SP cells show long-term repopulating capabilities

An important characteristic of a CSC is that it gives rise to a fully heterogenous population with cells that are capable of self-renewal as well as more differentiated progeny, by asymmetric division. In order to investigate this property, SP cells from the Raji cell line were isolated and cultured for a three week period. At days 6, 12 and 18 cells were re-analysed for the SP phenotype as previously described. Over time there was a decrease in the SP fraction reducing from 100% at day 0 to 30% by day 6 and 2.47% of the population by day 18 (Figure 5.7). Not only do these data show that the SP can give rise to both SP and MP cells, it also highlights the tendency of the cell line to maintain a tightly controlled population equilibrium. This phenotypic equilibrium in cell state proportions has been demonstrated in other SP cell line studies and has also been extensively studied with respect to cell surface protein expression[214,489]. These studies highlight the importance of maintenance of a steady state within many cell lines even though they are considered autonomous and therefore independent of extracellular factors.

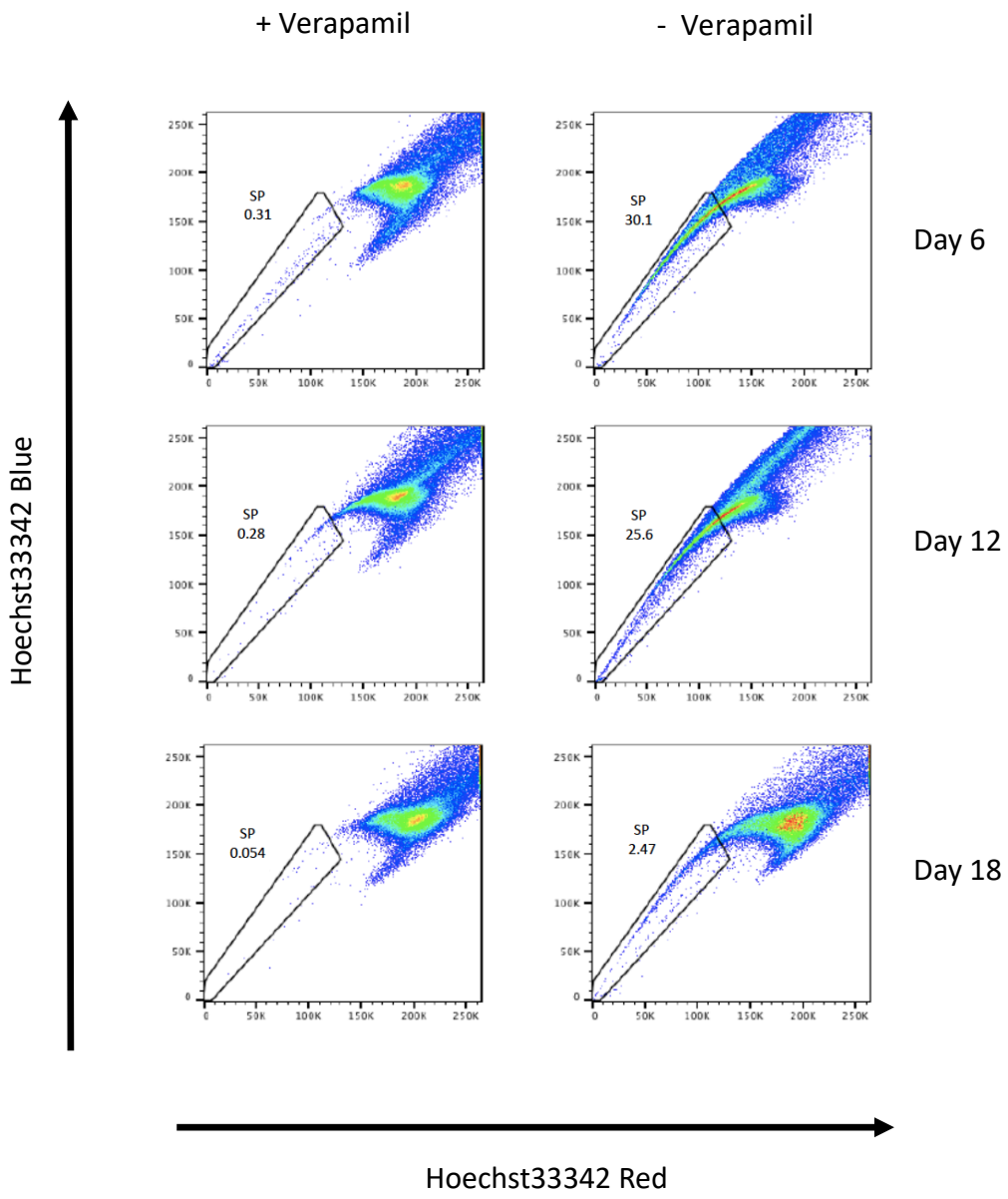


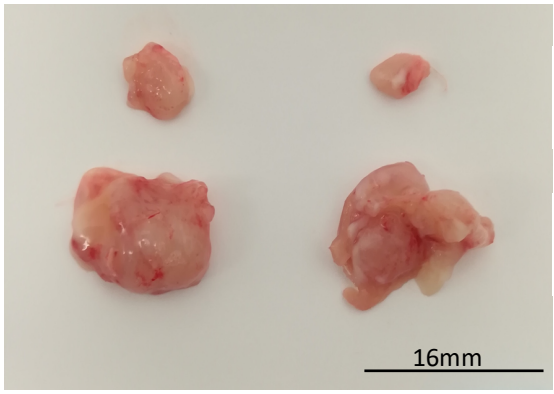
Figure 5.7. SP cells give rise to both SP and MP populations

Isolated SP cells from the Raji cell line were grown in vitro and the proportion of SP cells reanalysed at 6 day intervals. Data are representative of 2 independent experiments.

5.2.6 SP cells purified from BL cell lines are more likely to produce large tumours in NSG mice than MP cells

Having examined the growth dynamics of the SP and MP fractions *in vitro*, their tumourigenic potential *in vivo* was assessed. NSG mice were injected SC with SP or MP Raji cells ($5X10^4$), suspended in matrigel. Tumour growth was monitored over the next 40 days and when the first tumour reached its maximum allowed size all mice were culled and tumours extracted. Both the SP and MP injected mice produced tumours although the MP tumours were significantly smaller in size ($p<0.005$; Figure 5.8 A and C).These data were replicated for all of the cell lines and in all cases either the MP did not produce a tumour (i.e. Bjab and DG-75) or the tumours were significantly smaller in volume than the SP-derived tumours (Figure 5.8 B and C). Of note, for Bjab and DG-75 cell lines, although the experiment was stopped when the first SP tumour reached its maximum allowed size, the MP-injected mouse which had not yet produced a tumour was not culled. In both cases, after 6 months, no tumour formed and the mouse was culled.

A.



Raji MP-derived tumours

Raji SP-derived tumours

B.

Cell Line	Tumours produced from SP-injected mice (n=2)	Tumours produced from MP-injected mice (n=2)
Raji	**2**	**2**
Bjab	**2**	**1**
DG-75	**2**	**1**
Ramos	**2**	**2**

C.

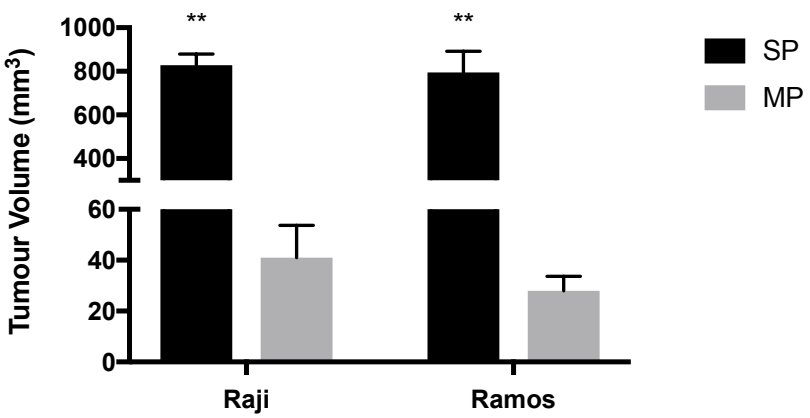


Figure 5.8. SP cells are more likely than MP cells to produce tumours in NSG mice.

*(A). SP cells give rise to significantly larger tumours than MP cells of the Raji cell line. (B) The tumour incidence in SP and MP-injected mice for each cell line. (C) Tumours produced from MP cells isolated from Raji and Ramos cell lines were significantly smaller in size than SP-derived tumours. **p<0.005*

5.2.7 SP cells are more resistant to chemotherapy than MP cells but are sensitive to PI3K/mTOR inhibition by Dactolisib

Given the important role SP cells have in driving tumour growth, efforts to target and eliminate them have important implications towards successful therapy. SP cells were isolated from each of the three cell lines before treatment with doxorubicin or methotrexate at the relevant IC_{50} for 72 hours. Cell viability in comparison to untreated control cells was determined using the cell titre blue assay (Figure 5.9). For all three cell lines, the SP had significantly increased viability compared to whole cell populations treated with doxorubicin (Figure 5.9A) or methotrexate (Figure 5.9B).

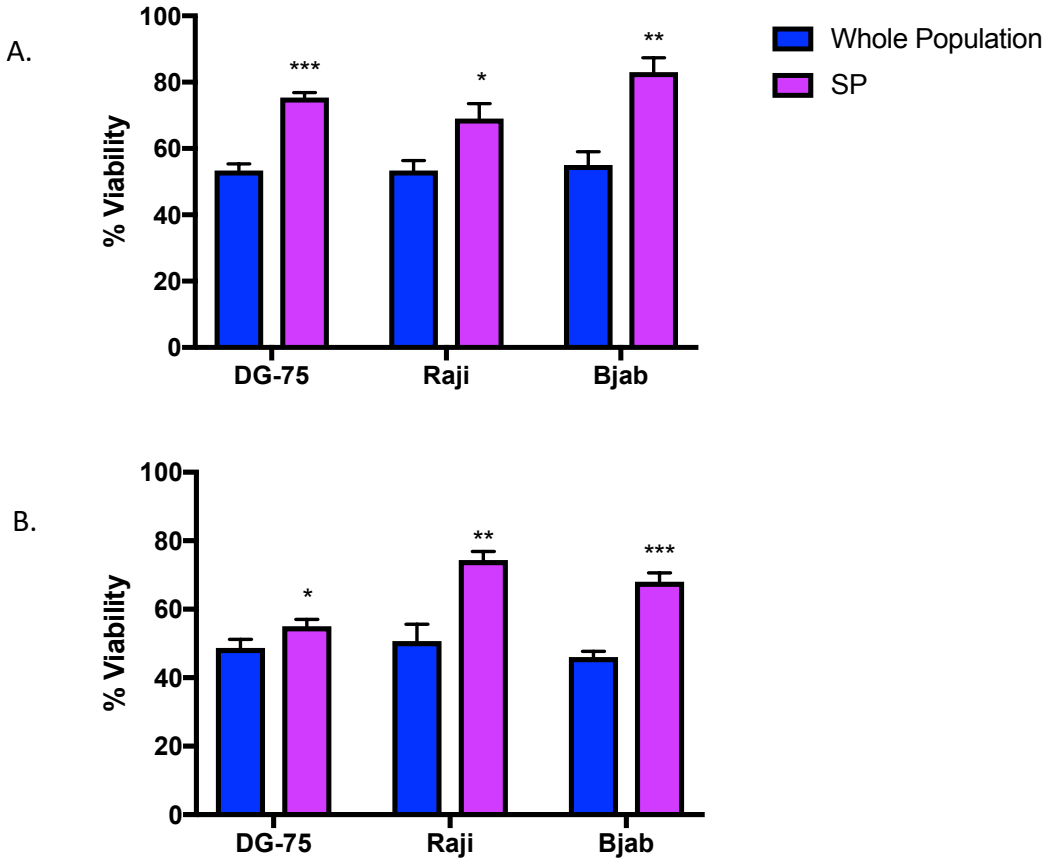
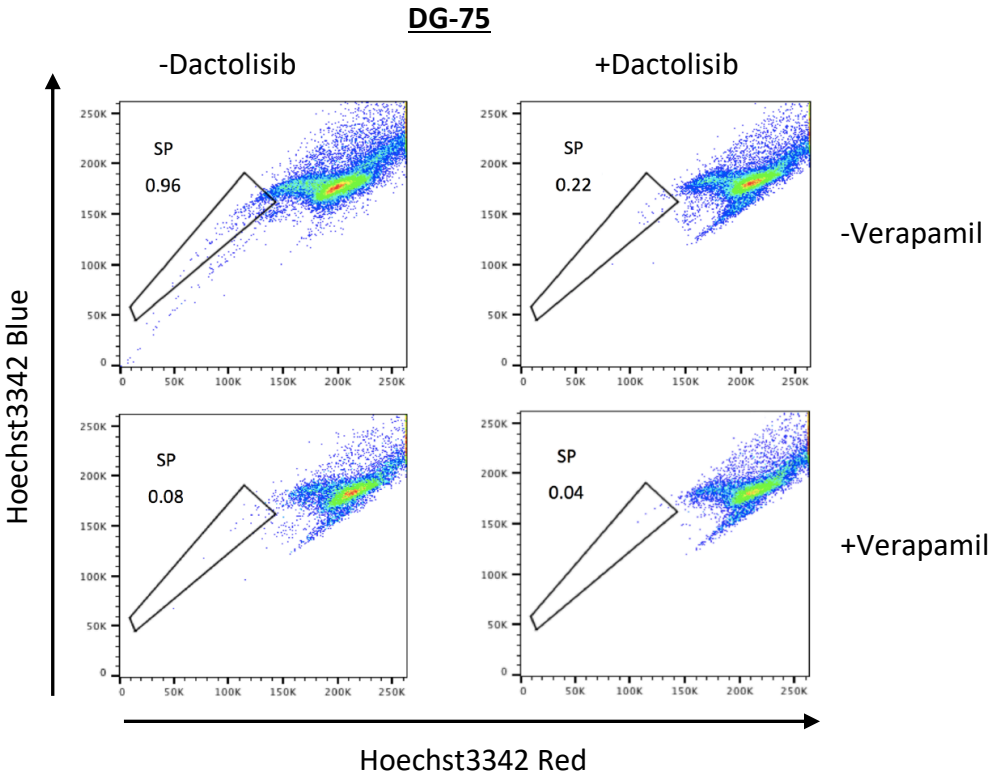


Figure 5.9. SP cells are more resistant to chemotherapy than the whole cell line population

*SP and whole cell populations were treated with Doxorubicin (A) or Methotrexate (B) at the relevant cell line IC_{50} for 72 hours. Cell viability was then determined using the CellTiter-Blue viability assay compared to untreated controls. Data are representative of the means and standard deviations of technical triplicates. *p<0.05, **p<0.005, ***p<0.001*

In order to determine the chemotherapeutic sensitivity of SP cells isolated from the cell lines, they were exposed to a variety of chemotherapeutic agents currently under investigation for the treatment of B cell malignancies. These included doxorubicin, methotrexate, vincristine, the MYC inhibitor JQ1, the BTK inhibitor Ibrutinib, the anti-CD20 antibody Rituximab and the PI3K/mTOR inhibitor Dactolisib. Only Dactolisib treatment resulted in a significantly reduced SP fraction compared to untreated controls in all three cell lines (Figure 5.10). Other drugs also had some impact on the overall SP percentage, including JQ1 which reduced the population in the DG-75 cell line. Furthermore, doxorubicin increased the SP proportion in all three treated cell lines and methotrexate enriched for the SP in two cell lines (Figure 5.11). Given the very significant results for Dactolisib, the reduced SP fraction in Dactolisib treated cell lines as well as the SP fraction from untreated cell lines were isolated by FACS. Further analysis identified a potential mechanism by which the PI3K/mTOR pathway regulates the SP. Flow cytometry analysis of untreated and treated SP cells highlighted decreased surface protein expression of ABCG2 in Dactolisib treated cells (Figure 5.12) and reduced mRNA levels of this transporter as well as ABCB1 were also validated by RT-PCR (Figure 5.12).



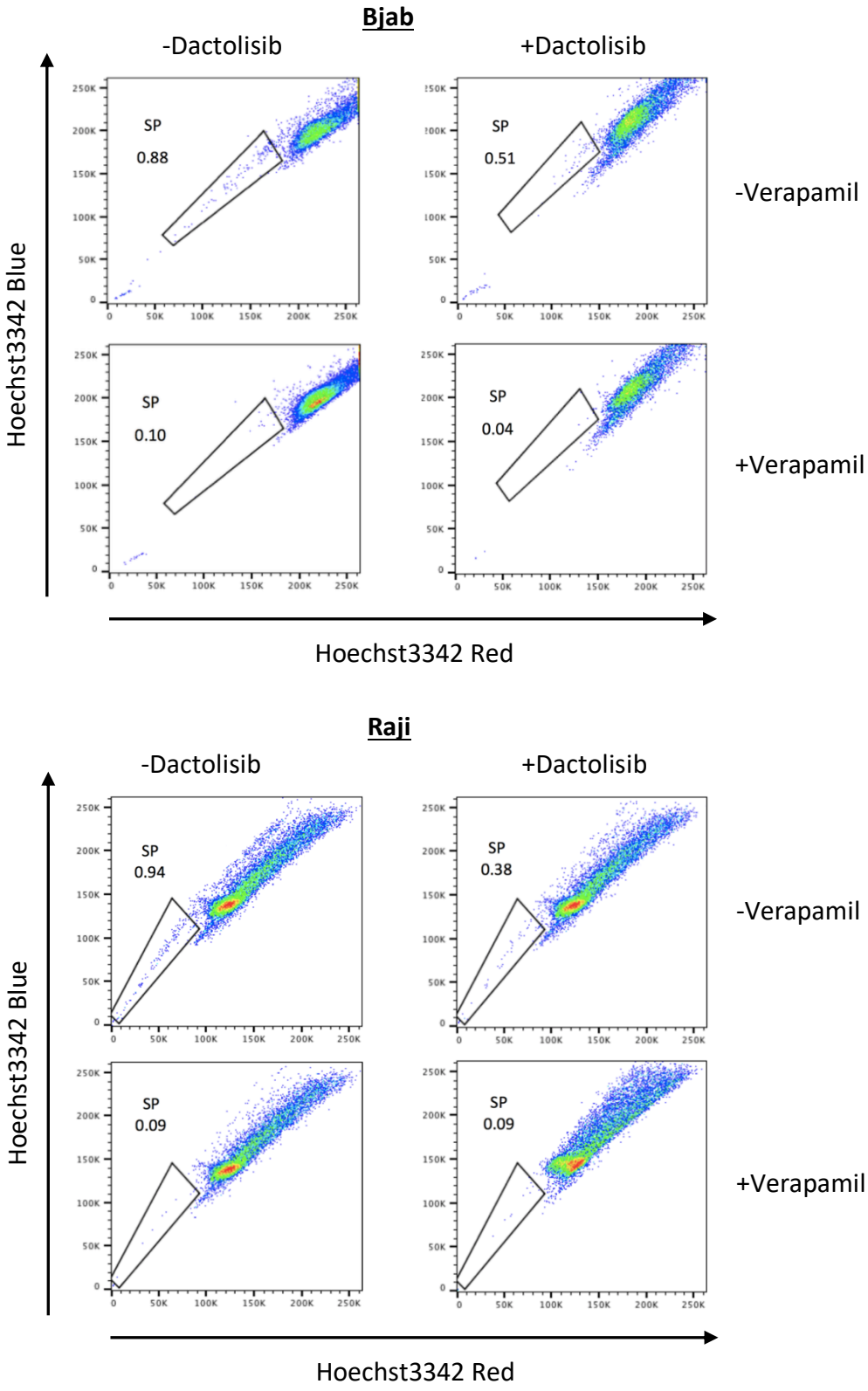


Figure 5.10. The SP is reduced in cell lines following Dactolisib treatment

Whole cell populations of the indicated cell lines were treated with Dactolisib at the relevant IC_{50} for 72 hrs. Live cells were then analysed for the presence of SP cells in the presence and absence of verapamil control.

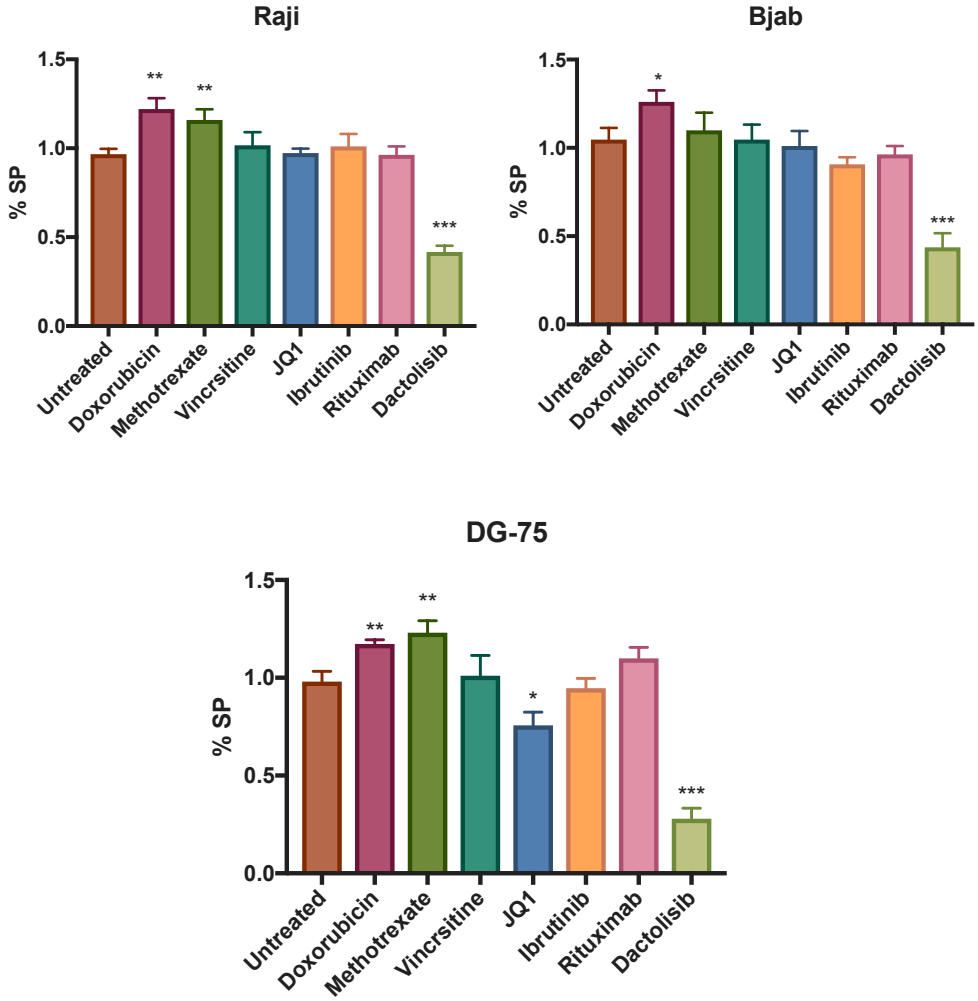


Figure 5.11. The proportion of SP cells in cell lines is reduced following Dactolisib treatment

*The indicated cell lines were treated with a range of chemotherapeutic agents as specified at the relevant IC_{50} for 72 hours before analysis for the presence of live SP cells. Data are representative of the means and standard deviations of technical triplicates. *p<0.05, **p<0.005, ***p<0.001*

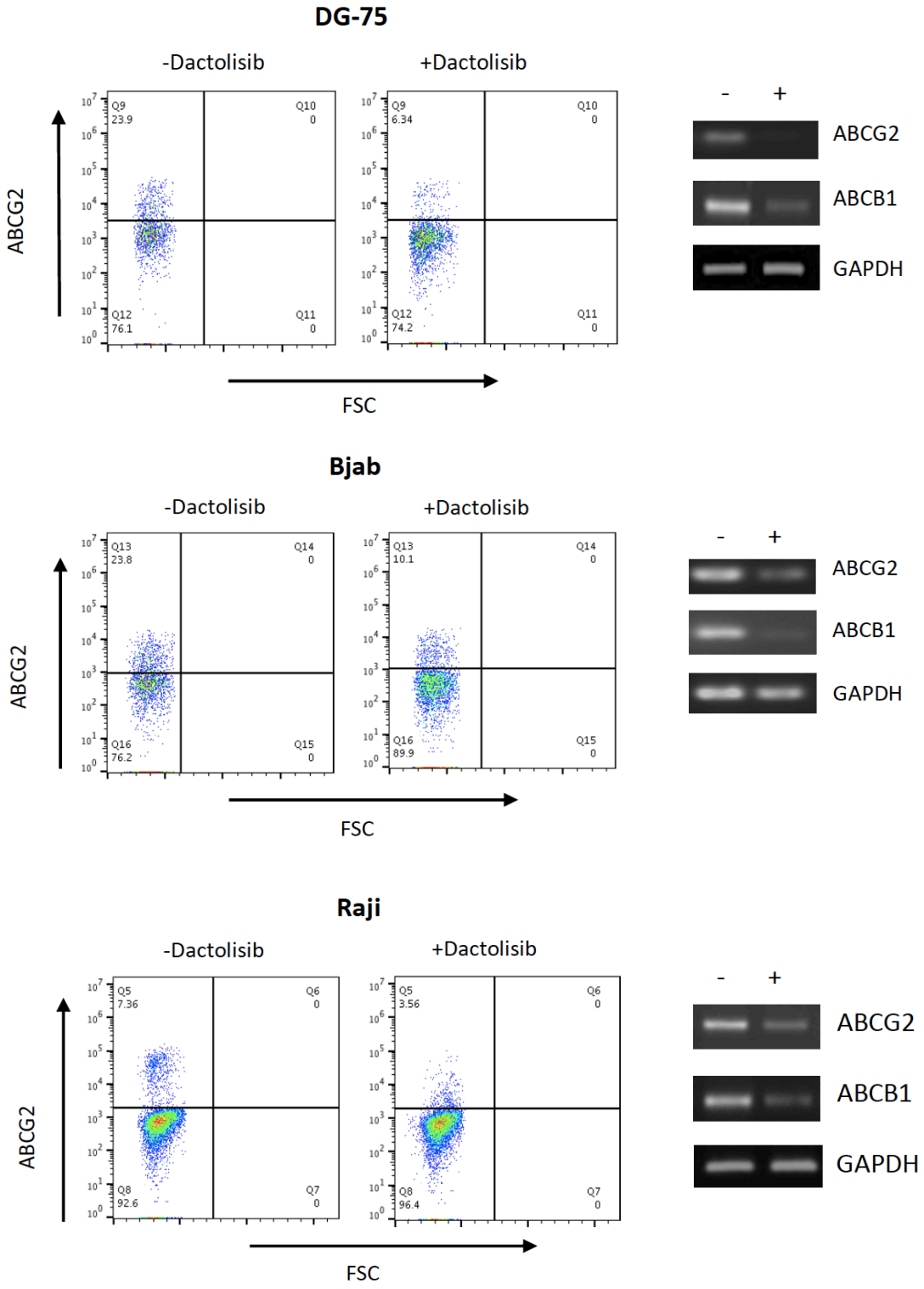


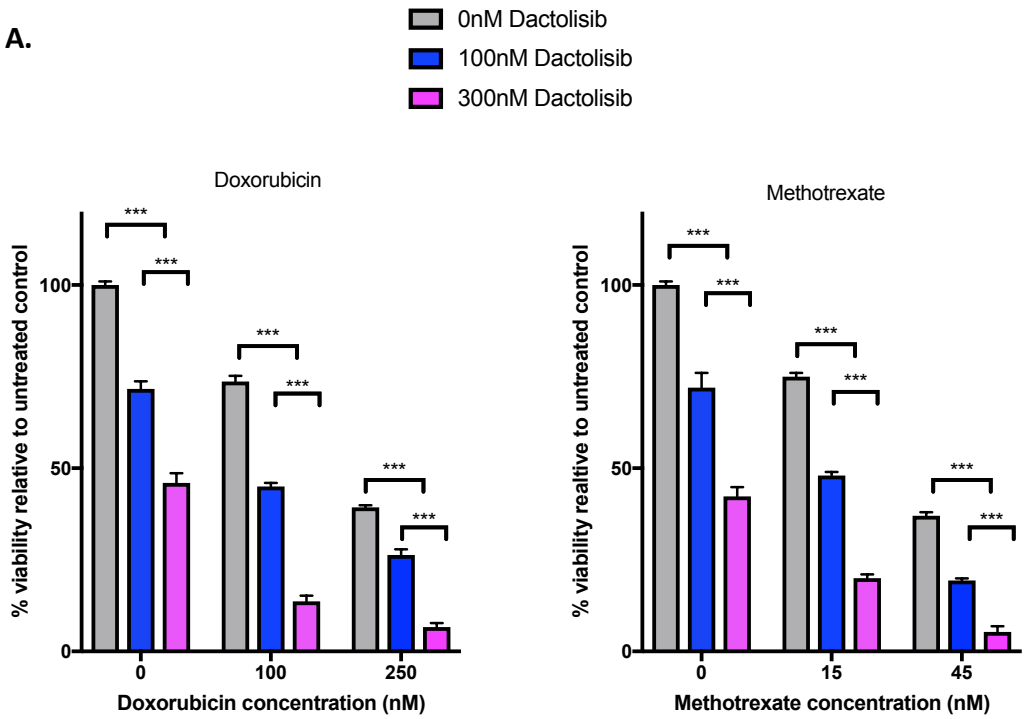
Figure 5.12. Treatment with Dactolisib reduces expression of ABC transporters in BL cell lines.

SP cells from both Dactolisib and control treated cell lines were sorted and labelled with ABCG2 antibody before analysis by flow cytometry. RNA was also extracted from the SP cells and RT-PCR carried out to detect transcript levels of ABCG2 and ABCB1.

5.2.8 Exposure to Dactolisib sensitises the Raji cell line to chemotherapy induced cell death

Given the efficacy of Dactolisib at targeting the SP, a population with increased tumourigenic potential and chemoresistance, we next investigated if combined treatment of the chemotherapeutics doxorubicin and methotrexate with this PI3K/mTOR inhibitor would induce synergy. Two concentrations of doxorubicin and two concentrations of methotrexate were each combined with two concentrations of Dactolisib (Figure 5.13A). At each combined concentration a synergistic effect was observed with the greatest synergy observed for the combination of 100nM doxorubicin with 300nM Dactolisib (Figure 5.13B).

A.



B.

Doxorubicin (nM)	Dactolisib (nM)	
	100	**300**
100	0.85	0.76
250	0.87	0.86
Methotrexate (nM)	**Dactolisib (nM)**	
	100	**300**
15	0.88	0.85
45	0.90	0.89

Figure 5.13. Dactolisib sensitizes the Raji cell line to doxorubicin and methotrexate

*(A.) The Raji cell line was treated with doxorubicin or methotrexate either alone or in combination with Dactolisib at the concentrations displayed. Cells were treated for 72 hours and viability determined using the CellTiter-Blue assay, relative to untreated control. Means and standard deviations of technical triplicates are shown. Representative of 2 biological replicates ***p<0.001. (B.) The Bliss Combination Index (BCI) for each combined treatment using the BCI formula, BCI<1; synergistic, BC1=1; additive BCI>; antagonistic.*

5.3 Discussion

The use of functional assays such as the SP phenotype to identify potential CSC populations has been a valuable resource and has contributed to the advancement of knowledge in the CSC area across numerous haematopoietic and solid malignancies. Indeed, cell lines have proven very useful as a resource to study the SP phenotype and this has been widely reported on[458]. Similar to investigations of the PDXs in chapter 4, analysis of cell lines identified expression of heterogenous surface proteins including ABC transporters and primitive markers (data not shown) and hence the presence of a SP was investigated. Following assay optimisation, the existence of a SP was verified in all 4 cell lines. The significant increase in the SP percentage after altering FBS content of culture media from 10% to 2% was an interesting finding. The influence of culture conditions and environment on the SP phenotype has been documented in other studies and the increase in the population when cells were cultured under conditions of serum starvation may imply that there was selective survival of SP cells or that cells transition to this phenotype under stressful conditions[356,484].

The limited proliferative capabilities of the MP cells compared to the SP in all four cell lines is a particularly interesting observation. Conflicting data exists regarding the proliferative potential and cell cycle dynamics of the SP. Some groups have assigned slow-proliferative capabilities to the SP[347,490] while others have reported a high proliferation rate for this population[491–493]. An explanation for these observed differences may be centred on the differentiation stage from which the SP arises. For example, if the SP originates from an early, slowly proliferating progenitor such as a tissue specific stem cell then this characteristic may be retained by the SP. However, if the SP arises from a cycling cell such as an activated B cell then the high proliferative potential may be conserved. The high proliferation rate of the SP in this study further suggests an origin of the SP in a mature B cell that has acquired features of stemness via an accumulation of genetic, epigenetic and environmental conditioning. The majority of published studies investigating SP cells focus on colony forming assays to highlight the growth advantage of the SP phenotype. This includes studies in renal cell carcinoma, small cell lung cancer and ovarian cancer[477,494,495]. Fewer studies examine the *in vitro* growth rates of SP and MP cells but those that do highlight a similar trend to our data,

with the MP population having limited proliferative capacity compared to SP. In Multiple Myeloma non-SP cells showed no proliferation until day 10 in culture and then showed significantly slower growth than SP cells[496]. In ALCL, after 48 hours in culture the majority of MP cells were unable to survive and proliferate and in the Neuroblastoma cell line SK-N-SH, SP cells expanded 500 fold *in vitro* over a two week period compared to only 2.9 fold for the MP population[430,345]. Of note, when MP populations in the BL cell lines analysed were cultured in transwell assays with SP cells, allowing for exchange of soluble factors, MP growth was recovered. The concept of a minority population of cells having such a large impact on the majority population is a difficult concept however, this self-maintained intra-clonal equilibrium of tumour cell populations has been well studied in recent years. Of particular interest, a pivotal study of DLBCL by Koch *et al*., showed that SP cells possessed a dual function in the propagation of the lymphoma bulk[230]. Besides generating progeny, the SP cells provided Wnt3a to surrounding non-SP cells. Clonogenicity was restricted to the SP phenotype when populations were separated although upon co-culturing GFP-labelled SP and non-labelled MP, the ability of MP cells to give rise to SP cells was observed. This transition was shown to be due to the transfer of Wnt3a from SP to MP cells via exosomes. This study supports the idea reviewed by Snijder *et al.,* that tumour population expansion does not represent stochastic growth from independent cells but rather a deterministic process, where fate decisions depend on the cellular context or the 'neighbourhood'[497]. The original idea that the CSC was a static, well-defined sub-population of cancer cells with invariable functional characteristics distinguishing them from cells of the tumour mass, has been replaced by the idea of CSC plasticity whereby the CSC phenotype is a transient state that many cells may acquire depending on cues provided by its microenvironment[207]. This highlights the importance of not just identifying the CSC population present at any one particular time but also of identifying the drivers of this phenotype.

In our study, exosomes isolated from SP cultured cells did not recover MP cell growth. However, exosome isolation is a difficult technique that requires significant optimisation. The concentration of isolated exosomes achieved was not very high, hence in light of the above study in DLBCL it may be worth further exploring a role for exosomes in driving MP cell growth. In particular, the proportion and characteristics of exosomes

released from SP and MP cells should be compared. However, this study did identify differential production of cytokines by the SP and whole cell populations in three cell lines. Bjab SP produced 3 cytokines – BAFF, CD40L and IL-10 at significantly higher levels whilst Raji SP cells also produced greater amounts of IL-10 and DG-75 SP cells showed higher production of IL-11 and BAFF. No cytokines were significant to all three cell lines although IL-10 and BAFF were both significant in two of three cell lines each. The impact of the microenvironment on tumour cells, in particular the complex network of stromal cells, signalling molecules, soluble factors and the extracellular matrix has seen an explosion of research in recent years. In particular, cytokines have emerged as important players in tumour promotion and progression and their role in the CSC phenotype is under investigation in many labs. An early study in the Bjab cell line highlighted the potentially important role of cytokine signalling for the cell line dynamics. When cultured under suboptimal conditions only a minor subclone was capable of long-term growth, supported by autocrine growth factors, and most notably the supernatant of this subclone stimulated growth of the parental Bjab line and showed IL-1-like activity[498]. IL-1 was not in our cytokine panel and should be investigated in the future, although this data supports our findings that individual cytokines may have a very significant impact on the growth dynamics of BL cell lines and investigation of each of the cytokines we have identified is warranted.

An example of a cytokine identified as important in promoting stemness is IL-6. It has also been shown to modify the differentiation state of cancer cells through upregulation of a stem-like phenotype in melanomas, breast and lung cancers[226–228]. Interestingly, IL-11 which was produced to a higher level by SP cells of the DG-75 cell line is a member of the IL-6 family and although less well studied, is known to activate at least three major pathways: JAK/STAT3, PI3K/mTOR and Ras/ERK pathways[499]. Furthermore, in triple negative breast cancer, a stem cell-like population is characterised by expression of the IL-11 receptor although whether the SP of the DG-75 cell line express the IL-11 receptor remains to be determined[500]. CSCs and TICs have been shown to secrete anti-inflammatory cytokines such as transforming TGF-β, IL-13 and IL-10 *in vitro*[501] and Xu *et al*., determined that hypoxia-induced secretion of IL-10 from adipose-derived mesenchymal stem cells promotes growth and CSC properties of the BL cell line BL2[502].

Given the increased production of IL-10 by Bjab and Raji SP cells, it is worth pursuing the role of this cytokine in the SP phenotype. Initially, expression of its receptor on SP and MP cells will be investigated although encouragingly, a recent genomic expression signature study in pBL highlighted upregulation of the IL-10 receptor in this large cohort of patients[62]. The identification of significantly higher BAFF production by the SP of two cell lines is another noteworthy result. BAFF is a member of the TNF family and as well as having important functions in normal B-cell development, has been shown to promote survival and proliferation of malignant B cells. Bolkun *et al.,* recently demonstrated that signalling by BAFF and another TNF family member, APRIL, plays an important role in AML pathogenesis and correlates with resistance to cytotoxic therapy[503]. Furthermore, in breast cancer both APRIL and BAFF were shown to increase cancer stem cell numbers, by inducing pluripotency genes such as KLF4 and NANOG[504]. A recent study generated a novel BAFF-Receptor antibody and showed its remarkable single-agent anti-tumour effects; chimeric antibodies bound with high affinity to multiple human malignant B-cell lines and induced potent ADCC against several subtypes of human lymphoma and leukaemia. This included primary tumours from patients who had relapsed after anti-CD20 therapy, as well as against Ibrutinib-resistant and Rituximab-insensitive CD20-deficient variant lymphomas[505]. The identification of differentially expressed cytokines between the MP and SP populations therefore provides evidence of potential factors driving the dynamic equilibrium in these cell lines. Further studies should include testing of a broader, more exhaustive panel of cytokines before utilising a matrix approach to determine if any combination of cytokines can recover the full proliferative capabilities of the MP cells. In addition, whether the receptors for these cytokines are expressed on SP and/or MP should be determined.

The SP fraction in our cell lines also showed significantly increased tumourigenic potential *in vivo*, similar to studies in PDXs. Targeting the SP fraction is extremely important not only due to its potential clinical relevance but will also help elucidate the underlying biology driving this phenotype. In this study, drugs that are readily available and under investigation to treat B cell malignancies were assessed for their effects on SP cells including traditional chemotherapeutic agents as well as the anti-CD20 antibody Rituximab, the MYC inhibitor JQ1, the BTK inhibitor Ibrutinib and a PI3K/mTOR inhibitor

Dactolisib. SP cells were shown in this chapter to have increased resistance to doxorubicin and methotrexate compared to the whole cell line and therefore the increased proportion of SP cells in the doxorubicin and methotrexate treated conditions is intuitive. Only Dactolisib had a significant negative impact on the survival of SP cells in each cell line. Schmitz *et al.,* have previously identified a number of genomic abnormalities in sporadic BL cases and cell lines. In particular, activation of the PI3K/AKT pathway has been implicated in BL lymphomagenesis and increased PI3K/AKT activation has been associated with worse outcomes in adults with aggressive B-NHL[39]. The PI3K-ð isoform inhibitor Idelalisib, has been approved for relapsed and refractory FL, CLL, and small lymphocytic lymphoma. Although these agents are extremely active, they achieve primarily partial responses as single agents and require indefinite administration and hence their use in combination with chemotherapy may be a better strategy[506]. Very recent pre-clinical studies targeting the PI3K/AKT/mTOR pathway in models of BL have also shown promising potential[122]. The role of this pathway in CSCs has also been intensively study. For example, a potent and selective dual inhibitor of mTORC1/2 and class I PI3-kinases was able to inhibit proliferation and survival of breast CSCs *in vivo* and markedly reduced their tumour-initiating ability in limiting dilution assays[507]. It was also shown in paediatric Medulloblastoma that the CSC population is more sensitive to PI3K targeting compared to the whole cancerous population and its non-stem cell counterpart[508]. Other studies have also shown evidence of the underlying role of PI3K or AKT signalling in the CSC with Bleau *et al.,* demonstrating that AKT, but not its downstream target mTOR, regulates ABCG2 activity in glioma tumour stem-like cells[276]. Furthermore, activated PI3K upregulated ABCG2 expression and increased the percentage of cancer stem-like cells in both AML and ALL[275]. This agrees with data presented here that highlights not only that the percentage of SP cells was diminished by targeting the PI3K/AKT/mTOR pathway but also that ABCG2 cell surface expression was reduced and both ABCG2 and ABCB1 mRNA levels decreased following Dactolisib treatment. This is a very interesting insight to the role this pathway may be playing in driving the SP phenotype. Other pathways such as the JNK1/c-jun signalling pathway in colon cancer was also reported to be involved in ABCG2-mediated multidrug resistance, hence control of ABCG2 expression may not be specific to the PI3K/AKT/mTOR and other

pathways in the SP cells should be investigated too, particularly as Dactolisib did not completely eliminate SP cells[509]. Going forward it is of course important to utilise more specific inhibitors of both PI3K and mTOR to verify which is responsible for regulating ABCG2 and ABCB1 expression. Furthermore, the important observation that inhibition of PI3K/mTOR may have synergism with chemotherapeutics should now be investigated with the PDX models to determine if Dactolisib is more effective either alone or in combination with chemotherapy, compared to chemotherapy alone at inhibiting SP cells *in vivo.*

Overall, this chapter identifies the presence of SP cells in each of the BL cell lines and highlights the *in vitro* and *in vivo* growth advantage of these over MP cells. Furthermore, cell to cell communication is mediated via soluble factors which are important for MP cell growth. Soluble factors that are potential candidates mediating MP cell growth include cytokines secreted at a higher level by the SP, but the significance of these remains to be determined in future studies. Finally, ABC transporter expression on SP cells is mediated by the PI3K/AKT/mTOR pathway and more importantly might represent a therapeutic target for BL. It may be possible to target the SP, potentially by modulating expression of drug efflux transporters, increasing the sensitivity of SP cells to chemotherapeutic agents that are substrates for these.

Chapter 6. Relapse can be mimicked in PDXs and may identify resistance mechanisms in paediatric Burkitt Lymphoma

6.1 Introduction

Survival rates in pB-NHL have increased dramatically in recent decades to almost 90% although prognosis in refractory and relapse cases remains dismal, with very few options to successfully salvage these patients[510,511]. Gain of chromosome 7 and deletion of 13q are the only abnormalities that have been associated with an adverse prognosis in the context of a clinical trial[65]. Gains of 13q have also been reported, including gain of the MIR17 host gene (MIR17HG) locus and are associated with a tendency for early relapse[69]. The rarity of relapse cases combined with the fact that re-biopsy at relapse is not a routine procedure has meant that genomic and expression analysis of relapse and refractory pBL has not been extensively conducted and beyond the mentioned copy number alterations, that identify poorer prognosis, little knowledge exists regarding the drivers of relapse in pBL.

Resistance to both chemotherapy and targeted therapy is a significant problem across the spectrum of B-NHLs. Several mechanisms of resistance have been proposed including pharmacokinetic and pharmacodynamic events along with microenvironmental and tumour intrinsic characteristics. At the tumour intrinsic level, a number of possible drivers of resistance include increased drug efflux or drug inactivation, evolution of mutants of the target protein, repair of drug-induced DNA damage, activation of alternative signalling cascades, as well as deregulation of apoptosis[512,513]. Common anti-apoptotic proteins implicated in relapse to chemotherapy in several tumours include the BCL2 family members while mutations in *TP53* leading to avoidance of apoptosis following DNA damage have frequently been reported including in lymphoma[514]. Increased drug efflux via upregulation of ABC transporters is a well-documented resistance mechanism and has been shown to be effective at elimination of several drugs critical to the pB-NHL regimen including doxorubicin, methotrexate and vincristine[151,512,515]. Of course, the more recent emerging knowledge regarding intra-tumour heterogeneity and CSCs has added further complexity to the study of relapse. Drug resistance and relapse is thought to be either intrinsic due to heterogeneity of the

tumour or acquired following ongoing genomic tumour evolution under the selective pressure of drug treatment, or indeed it can be both. In the former, a pre-existing minor sub-clone may possess characteristics that confer resistance (e.g. stemness) and provide a survival advantage over the major non-resistant tumour population. In acquired resistance, genomic evolution occurs due to both selective pressure applied by the drug as well as genomic/genetic instability caused by alkylating and DNA intercalating agents present in chemotherapy regimens[512].

Studies of sequential biopsy samples in DLBCL and FL have provided insight into the different mechanisms of relapse and progression[516,517]. Relapse can evolve directly from a sub-clone present at diagnosis with the potential for acquisition of additional mutations. Alternatively, relapse may occur via early divergent evolution from an earlier tumour precursor shared by both the primary tumour and the relapse[518]. Both models have been identified in DLBCL and FL and it has been suggested that these alternative models have important implications for therapy. Less data exists in relation to BL, although sequential karyotyping at diagnosis and progression/relapse of paediatric samples in one study highlighted that for the majority of cases, aberrations present in a sub-clone at diagnosis were also present in the relapse sample along with additional karyotypic complexity[519]. A recent case report described whole exome sequencing data of matched presentation and relapse biopsies from a nine year old patient in which relapse occurred after four months and salvage treatment failed[520]. Genomic analysis determined that relapsed BL evolved from one of the sub-clones observed at the initial phase and was accompanied by additional mutations including chromatin remodelling genes, suggesting chromatin regulation might be associated with disease progression. These studies highlight the importance of understanding the heterogeneous clones within pBL at presentation to better predict and treat relapse cases.

Extensive studies have highlighted the role of CSCs in drug resistance and tumour relapse. The characteristics of CSCs or TICs has varied across cancer types as well as their frequency, although it has been demonstrated that these malignant cells with stemness characteristics are extremely tumorigenic and are resistant to conventional chemotherapies and radiation. Such highly tumourigenic and drug-resistant cells are

therefore likely to survive chemotherapy or radiotherapy and ultimately be responsible for relapse[415,521–524]. However, cells with stemness characteristics are difficult to target owing to activation of pro-survival and antiapoptotic pathways, overexpression of drug efflux pumps and increased DNA repair capacity[521–526]. Several therapeutic approaches to target stemness have been suggested including targeting of the stemness-associated signalling cascades and the microenvironment as well as targeting ABC transporters and surface markers that identify cells with stemness (reviewed in ref 527). In studies utilising the SP assay to identify the CSC, this population was shown to contribute to relapse. In cell lines derived from relapsed neuroblastoma the SP was found not only to be significantly increased in size compared to the pre-treatment cell lines, but the SP cells also displayed increased proliferative ability and colony-forming efficiency, highlighting that they were not efficiently targeted by the conventional chemotherapy used in treatment[528]. The SP was also shown to be responsible for chemoresistance in patient-derived melanoma tissues and interestingly displayed unique resistance mechanisms for different drugs. ABC transporters mediated resistance against paclitaxel while SP resistance to temozolomide was mediated by IL-8[529]. The concept that the drug-resistant CSC population can have different driving mechanisms is a daunting one and within the PDX models in this thesis, various pro-survival and anti-apoptotic pathways were implicated in different patient samples. Efforts to target stemness to counteract relapse and metastasis are gaining pace and indeed the first in class cancer stemness inhibitor, targeting STAT3, has now reached stage III clinical trials highlighting the potential in this area[530]. In this chapter, the impact of chemotherapy on the heterogenous PDX tumours *in vivo* was evaluated to determine potential drivers of resistance in pBL samples.

Aims of this chapter

- Investigate the sensitivity of PDX SP cells to chemotherapy *in vivo*
- Mimic tumour relapse *in vivo* by administering chemotherapy to mice
- Investigate SP cells in relapse PDXs
- Carry out RNA-seq analysis to compare pre-treatment and relapse tumours.

6.2 Results

6.2.1 PDXs show heterogeneity with respect to drug response *in vivo*

Inherent inter-patient heterogeneity may lead to varying levels of response to treatment. In order to investigate this in the PDX models, mice bearing PDX1 and PDX4 tumours were treated with chemotherapy. Initial doses of chemotherapy were determined based on doses of doxorubicin, methotrexate and vincristine utilised in the induction stage of the Inter-B-NHL Clinical Trial[531]. These doses were converted to animal doses using the conversion factor of Freireich *et al.,* as recommended by the FDA[532]. The human equivalent doses of doxorubicin, methotrexate and vincristine – 60mg/m^2, 3000mg/m^2, 2mg/m^2, respectively equated to 30mg/kg, 1500mg/kg and 1mg/kg in animal doses. However, analysis of the literature for lethal doses of these drugs for mice highlighted that these quantities of methotrexate and doxorubicin were far in excess of what could be tolerated by mice. Hence, lower doses of 3mg/kg doxorubicin and 15mg/kg methotrexate along with 0.25mg/kg vincristine were chosen for initial analysis of the sensitivity of PDX tumours to treatment. Tumour cells ($5X10^5$) from PDX1 and PDX4 were injected sub-cutaneously into two NSG mice each and when tumours reached 400mm^3 chemotherapy was administered every third day and tumour volume carefully monitored. An interesting observation was the fact that PDX4 tumours showed very little response to this dose of chemotherapy and continued to increase in volume after two treatments, while PDX1 tumours were highly responsive. Due to the lack of tumour-responsiveness for PDX4, the doses of each of the 3 chemotherapy drugs were increased by 25% from the third dosing onwards, resulting in a reduction in tumour volumes (Figure 6.1).

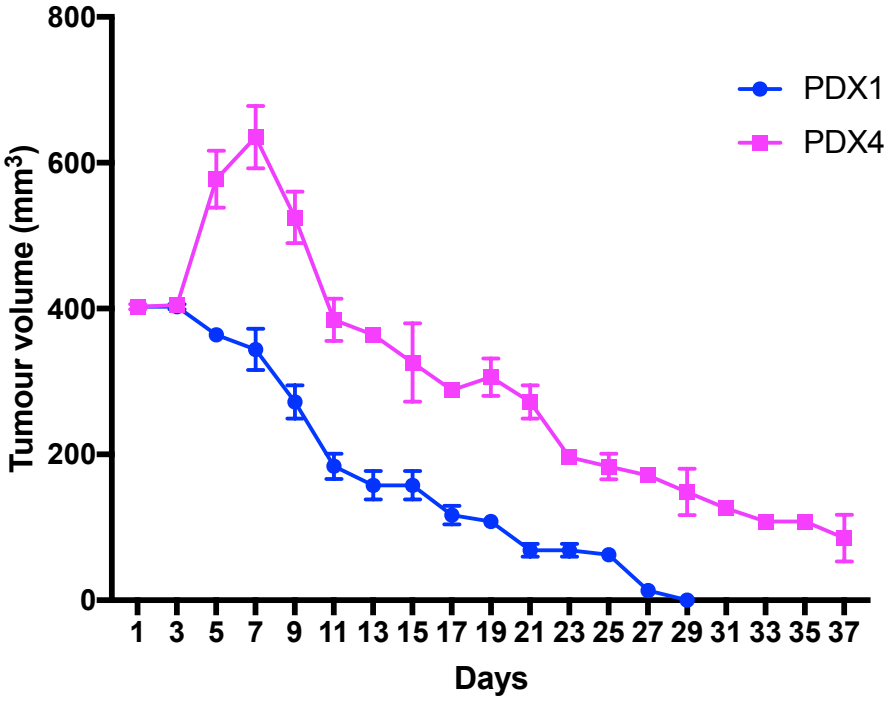


Figure 6.1. PDX4 is more resistant to chemotherapy *in vivo* than PDX1

PDX1 and PDX4 tumours were treated in vivo with multiagent chemotherapy at 72 hour intervals beginning when the tumour volume reached 400mm^3. Tumours size was monitored daily. Drug concentrations; 3mg/kg doxorubicin, 15mg/kg methotrexate, 0.25mg/kg vincristine. Increased to 3.75mg/kg doxorubicin, 18.75mg/kg methotrexate, 0.3125mg/kg vincristine for PDX4 after the second dosing. Means and standard deviations of duplicates shown for two mice carrying each of the PDX.

6.2.2 SP cells are more resistant to chemotherapy *in vivo*

After identifying effective chemotherapy concentrations for PDX1, SP cells from untreated tumour were sorted by FACs and whole population cells were also passed through the sorter. Cells were counted in triplicate to ensure accuracy and $5X10^5$ SP or whole population cells were injected into four mice each. When a tumour reached $400mm^3$ in volume, doxorubicin, methotrexate and vincristine were administered every three days for a total of five doses. The animals were culled 48 hours after the last dosing. As all eight tumours did not reach $400mm^3$ on the same day, treatment starting points were staggered over two days but each mouse received the identical regimen. Upon completion of the dosing, animals were culled and tumours carefully measured to determine both volume and mass. Chemotherapy resulted in an approximate 10% decrease in overall body mass for all mice. SP-derived tumours show significantly increased resistance to chemotherapy compared to whole cell-derived tumours (Figure 6.2A-B). The impact of chemotherapy on animal health, indicated by weight loss, was consistent across all mice in both groups of the study (Figure 6.2C).

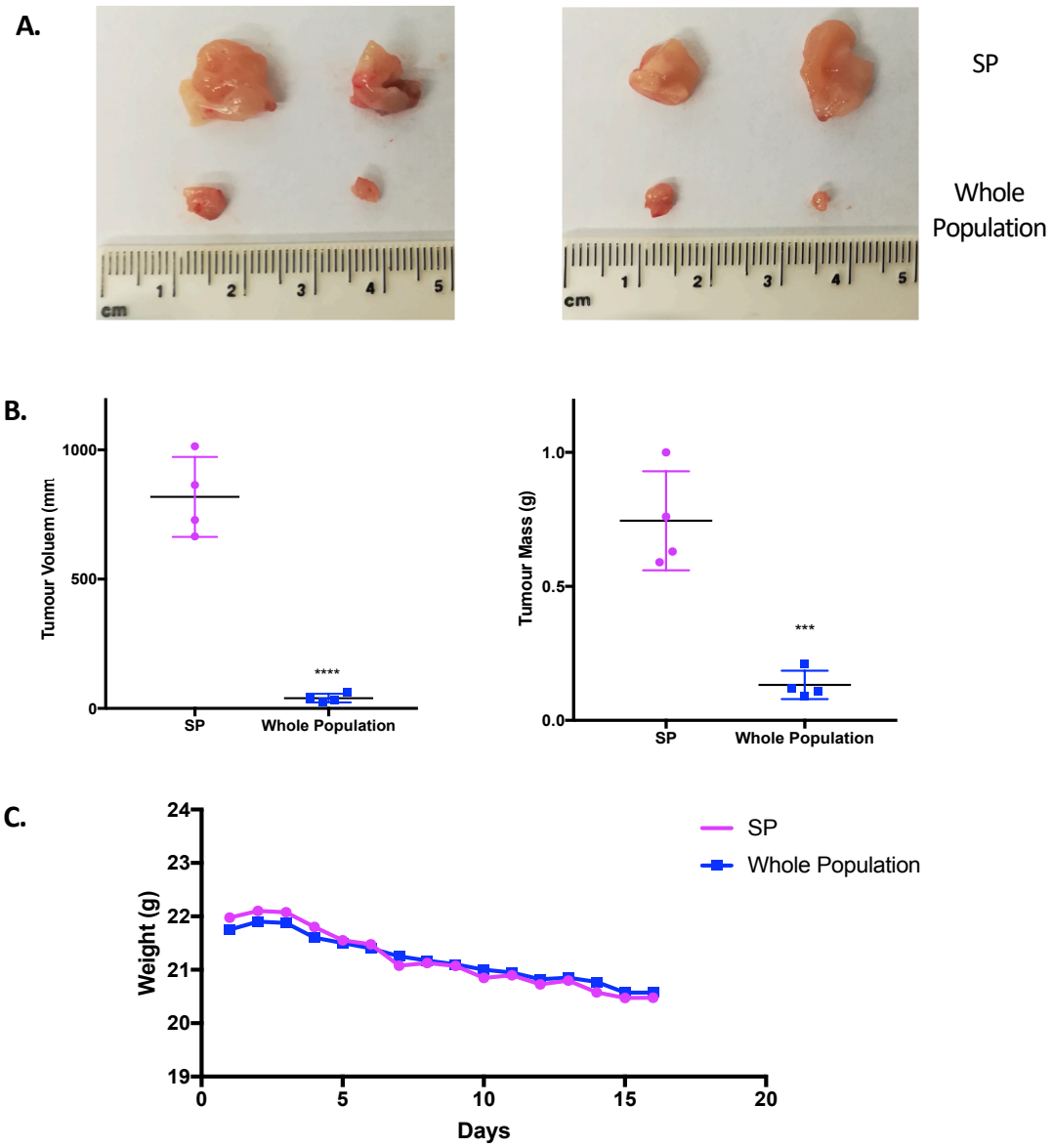


Figure 6.2. SP-derived tumours from PDX1 are more resistant to chemotherapy than whole population-derived tumours

*SP and whole population cells (5X10^5 of each) were injected into 4 NSG mice. Chemotherapy was administered every 72 hours starting when each tumour reached 400mm^3 in volume. Tumours were extracted 48 hours after the final dosing. (A.) The volume of all tumours were measured following extraction. (B.) Volume and mass difference between SP and whole population tumours. Means and standard deviations of four replicates shown *** p<0.001. (C.) Animal mass was monitored daily and means of the animal mass in each group displayed.*

6.2.3 PDXs can be used to mimic relapse *in vivo*

In order to determine potential mechanisms of relapse and resistance in pBL, efforts were made to induce relapse *in vivo* for all five PDXs following multi-agent chemotherapy. Following initial experiments to determine appropriate drug concentrations for each tumour, $5X10^5$ cells from each of the five PDX samples were injected into two mice each. When tumours reached $400mm^3$ one PDX-bearing mouse was culled for each of the five PDX tumours and the other mouse was administered chemotherapy by IP injection every three days (Figure 6.3). In all cases, tumour volume decreased with successive dosing until the tumour was deemed immeasurable. Tumour volume was measured every day independently by two individuals to ensure accuracy. The period of 'remission' ranged from 14 to 20 days in four PDXs however PDX4, which was the most chemo-resistant tumour and received the highest dosage of drugs, was only deemed immeasurable and in 'remission' for seven days. Apart from PDX1, the other tumours were extracted before they regrew to full size for reasons of animal welfare. Despite this, each tumour was a minimum of 7mm in diameter upon extraction, giving a plentiful cell population after tumour disaggregation for further analysis of the SP and RNA-seq. Growth dynamics of each PDX over the treatment course are displayed in Figure 6.4. Chemotherapy administration began at a volume of $400m^3$ with PDX 1,3 and 5 responding immediately. PDX 2 and 4 showed further growth following the first dose of chemotherapy but then began to reduce in size.

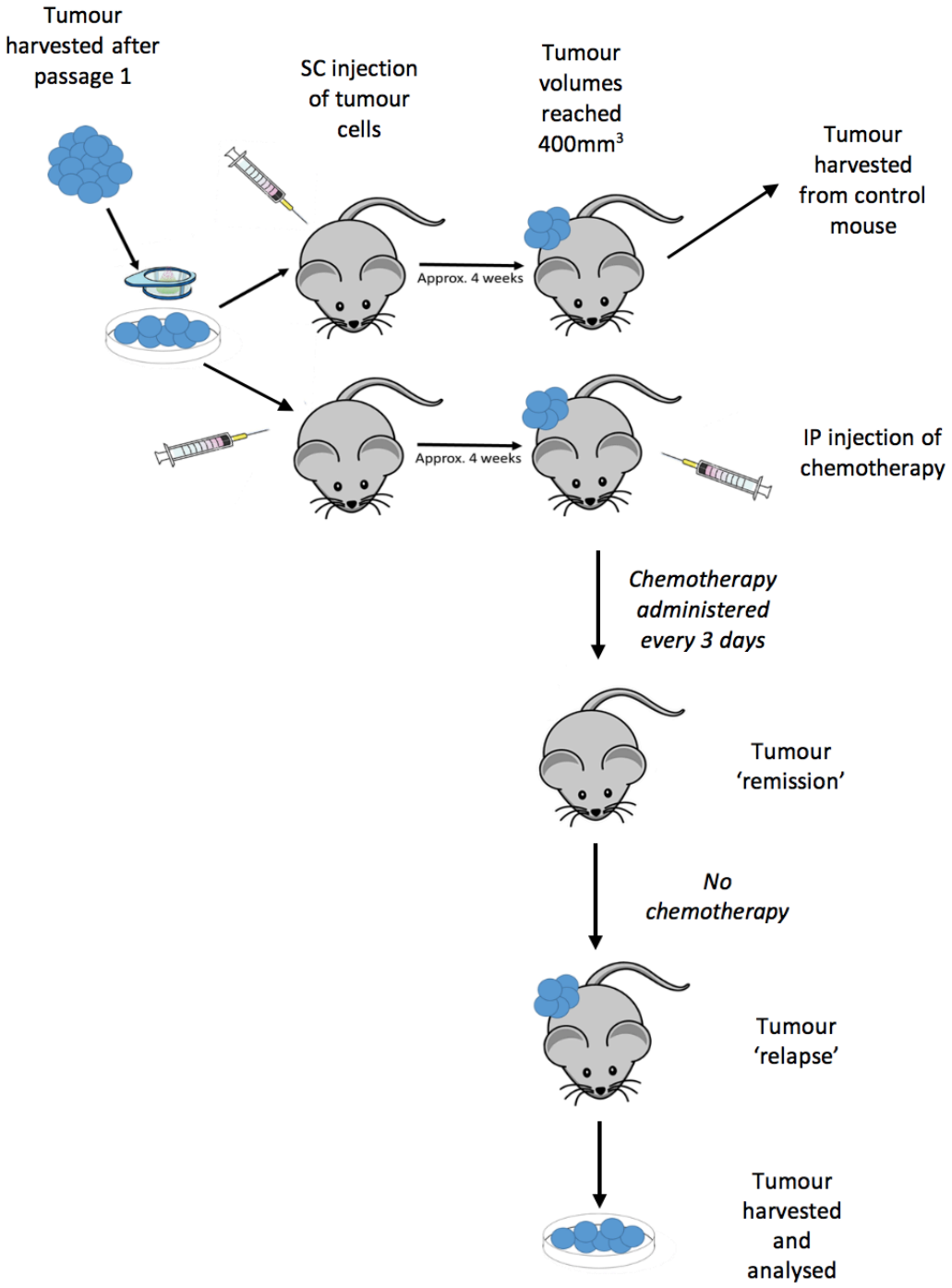
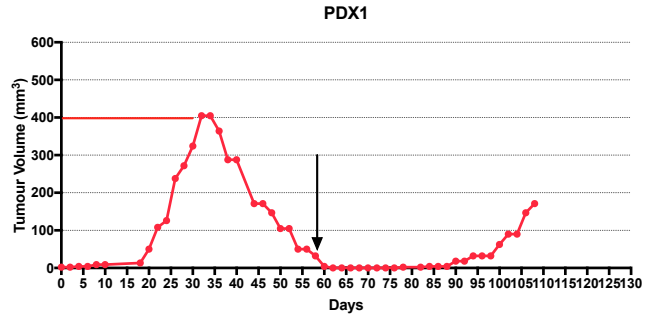
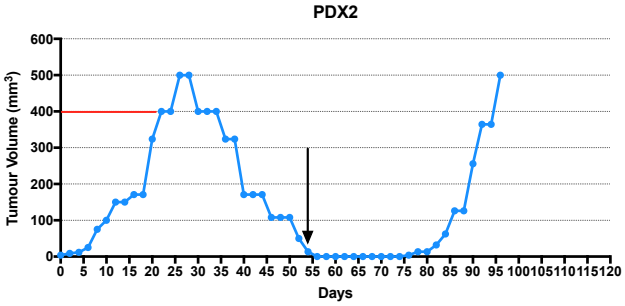
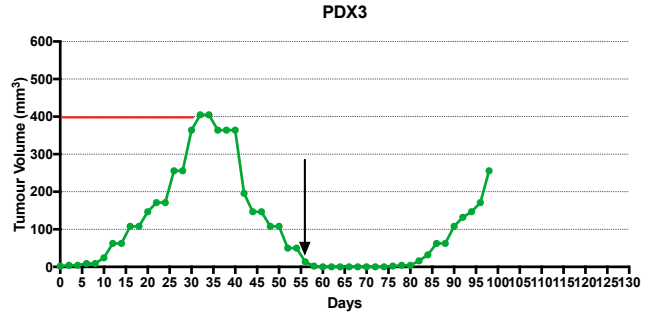


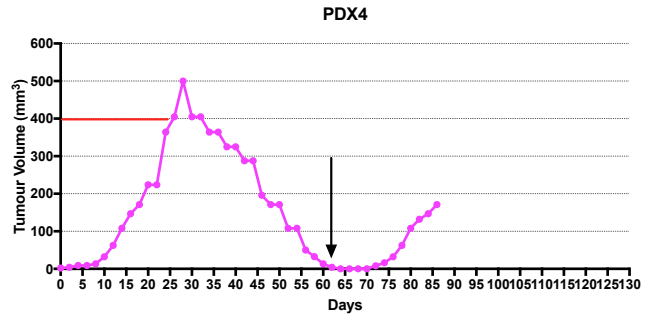
Figure 6.3. Schematic of *in vivo* PDX chemotherapy treatment protocol

Following passage 1, tumour cells from each PDX were injected sub-cutaneously (SC) into 2 mice each. When tumours reached 400mm^3 one PDX-bearing mouse was culled and the other mouse was administered chemotherapy by intraperitoneal (IP) injection every three days. Tumour volume was carefully monitored daily and treatment was ceased when the tumour was immeasurable. Following re-growth of the tumour, cells were harvested for downstream analysis.


PDX1
Tumour Volume (mm³)
Days


PDX2
Tumour Volume (mm³)
Days


PDX3
Tumour Volume (mm³)
Days


PDX4
Tumour Volume (mm³)
Days

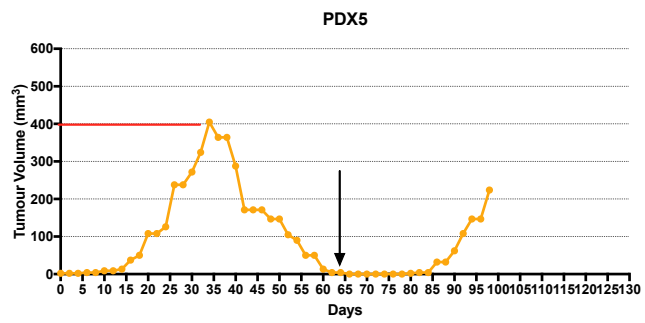
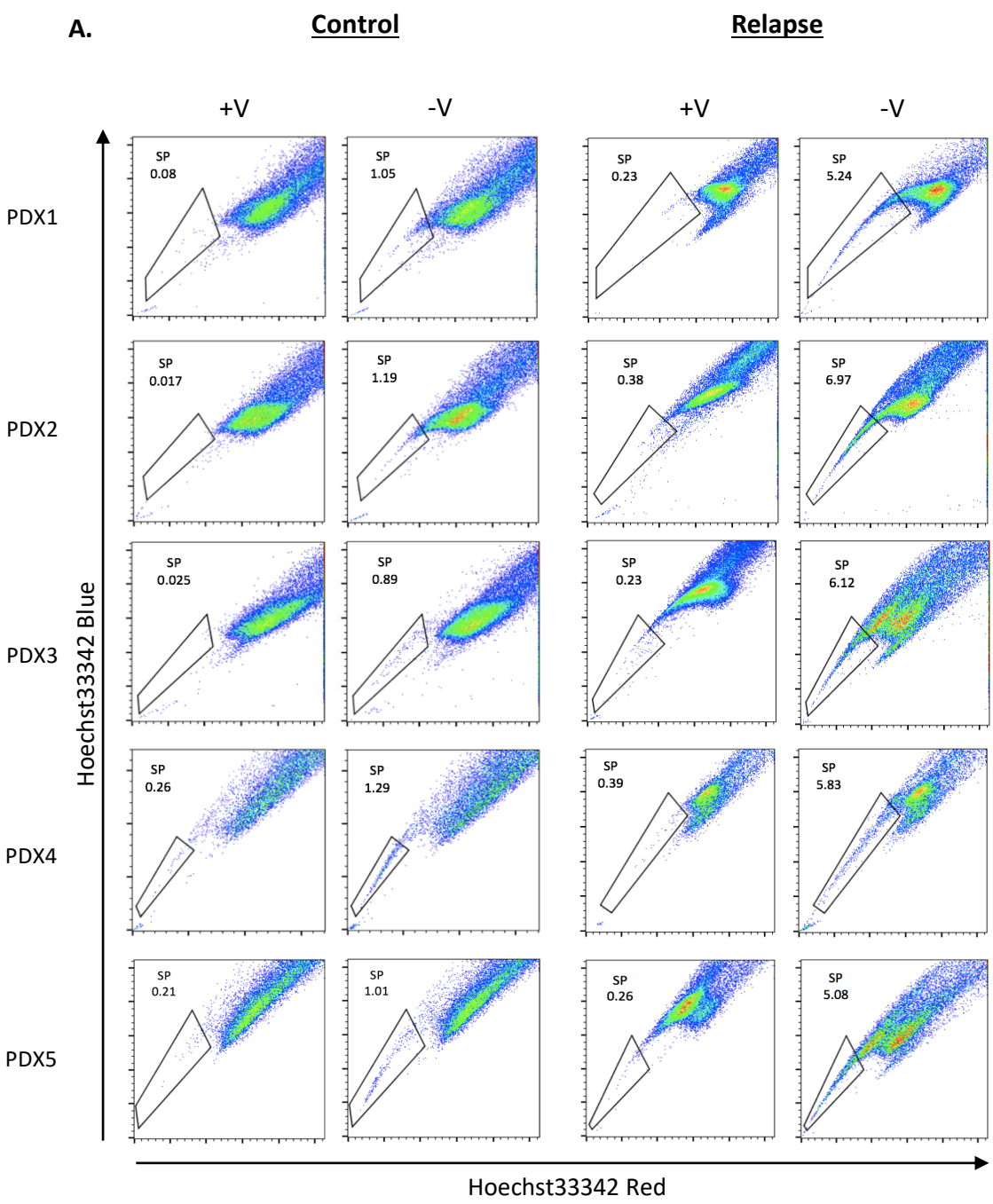

PDX5
Tumour Volume (mm³)
Days

Figure 6.4. Relapse can be induced in PDXs following successive multi-agent chemotherapy

Chemotherapy was administered IP to mice when PDX tumours reached 400mm^3 in volume (starting point denoted by the red line). Treatment was repeated every 72 hours and ceased when the tumour was immeasurable. Last chemotherapy administration denoted by black arrow. Graphs show tumour volume measurements from every second day. Tumour volume calculated using the modified ellipsoid formula 1/2(Length × Width^2).

6.2.4 The side population is increased in PDXs following regrowth after chemotherapy

Evidence from chapter 3 showed that SP cells have increased expression of ABC transporters and anti-apoptotic programs while this chapter highlighted the significantly increased resistance of SP-derived tumours to chemotherapy *in vivo,* compared to whole population tumours. In order to elucidate if the SP has a role in relapse of BL following recurrence of PDX tumours in mice, the tumour was extracted and disaggregated into single cell suspension and SP analysis carried out as previously outlined. In all five 'relapsed' PDX tumours, the SP is significantly increased in size compared to untreated controls (Figure 6.5 A and B).


A.
Control
Relapse
+V
-V
+V
-V
PDX1
PDX2
PDX3
PDX4
PDX5
SP 0.08
SP 1.05
SP 0.23
SP 5.24
SP 0.017
SP 1.19
SP 0.38
SP 6.97
SP 0.025
SP 0.89
SP 0.23
SP 6.12
SP 0.26
SP 1.29
SP 0.39
SP 5.83
SP 0.21
SP 1.01
SP 0.26
SP 5.08
Hoechst33342 Blue
Hoechst33342 Red

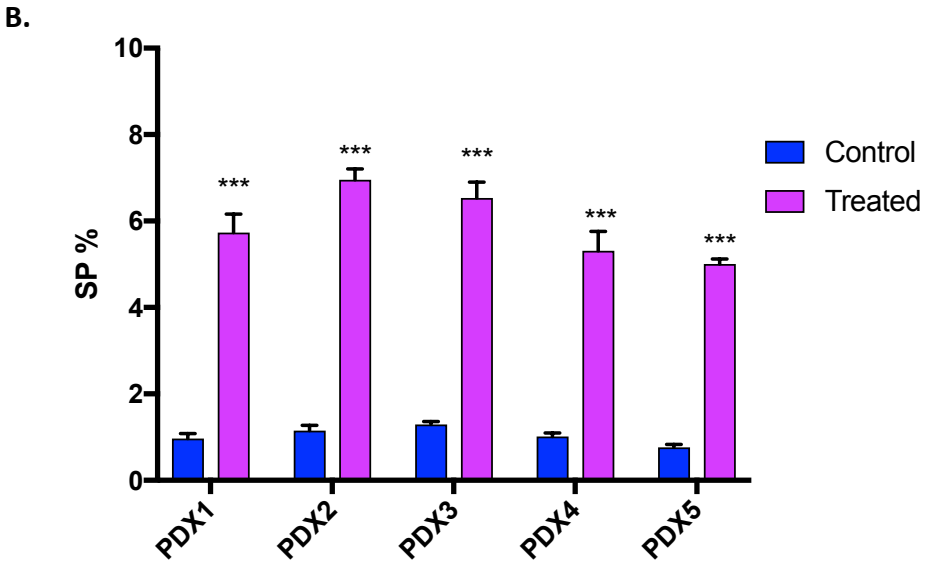


Figure 6.5. The SP is increased in PDXs following treatment with chemotherapy *in vivo*

*Control, untreated tumours were extracted when they reached 400mm^3 in volume and SP analysis performed. Treatment arm tumours received chemotherapy, were extracted when they re-grew and SP analysis performed. (A) Representative SP graphs for untreated and treated tumours along with verapamil (V) controls are shown for each PDX. (B) Means and standard deviations of technical triplicates are shown for each PDX. ***p<0.001*

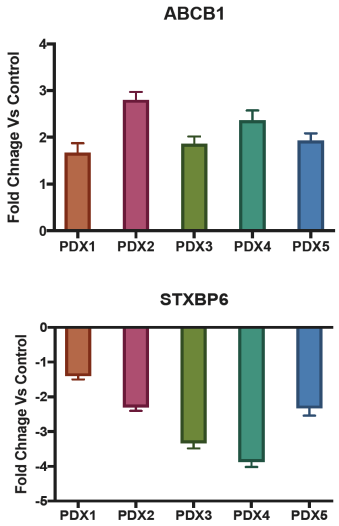
6.2.5 Differential expression profiles exist between relapsed and untreated PDXs

The increase in SP cells in each PDX after treatment, suggests their important role in relapse and resistance, although this population still makes up a small minority of cells in the treated tumours. In order to obtain a more global overview of expression changes induced by chemotherapy, RNA-Seq was carried out on the five matched pairs of untreated versus relapse whole population samples. Using a significance cut-off fold change > 1.5 and false discovery rate (FDR)-adjusted *p* value of 0.05, 23 differentially expressed genes were identified – 13 upregulated and 10 downregulated (Figure 6.6), suggesting that only a small number of genes are involved in tumour recurrence in all PDXs. Alternatively, due to the fact that our biological replicates are from five different patients there may be inter-patient as well as intra-patient heterogeneity, which together with a limited sample size, reduces statistical power. Despite the relatively small number of differentially expressed genes, functional annotation analysis of the hits using the molecular signatures database (MsigDB)[533] analysis tool identified gene ontology (GO) functions of up and down regulated genes. The upregulated genes were enriched in categories associated with cell cycle and response to stimuli, while the down regulated genes were enriched in categories involved in regulation of apoptosis and cell death as well as protein and cellular localisation (Figure 6.7).

A.

Upregulated genes	Fold Change	*p* value	FDR
MALAT1	3.11	$4.42X10^{07}$	$6.23X10^{04}$
COL1A1	2.93	$2.41X10^{06}$	$7.24X10^{04}$
CDC25B	2.42	$2.90X10^{06}$	$8.27X10^{03}$
ABCB1	2.41	$5.63X10^{06}$	$9.02X10^{03}$
HDAC10	2.28	$1.25X10^{05}$	$1.05X10^{02}$
CDK10	2.16	$2.21X10^{05}$	$9.63X10^{03}$
ADAM8	2.02	$4.72X10^{05}$	$9.98X10^{03}$
STAT5B	1.98	$5.89X10^{05}$	$1.43X10^{02}$
NFATC4	1.81	$7.02X10^{05}$	$2.91X10^{02}$
CDK11A	1.80	$1.094X10^{04}$	$3.35X10^{02}$
FANCD2	1.74	$1.34X10^{04}$	$3.67X10^{02}$
MAD2L2	1.66	$1.61X10^{04}$	$3.80X10^{02}$
SLC25A27	1.53	$1.67X10^{04}$	$3.83X10^{02}$
Downregulated genes	**Fold Change**	***p* value**	**FDR**
STXBP6	-2.94	$1.09X10^{07}$	$3.08X10^{04}$
LIMA1	-2.74	$3.89X10^{07}$	$4.78X10^{04}$
CD9	-2.52	$6.01X10^{07}$	$6.89X10^{04}$
CFL1	-2.50	$1.29X10^{06}$	$1.06X10^{03}$
CARD8	-2.39	$4.56X10^{06}$	$2.45X10^{03}$
PTPRCAP	-2.22	$6.78X10^{06}$	$4.15X10^{03}$
BID	-2.06	$7.93X10^{06}$	$1.19X10^{02}$
TGFB1	-1.79	$2.09X10^{05}$	$2.22X10^{02}$
DBN1	-1.69	$5.12X10^{05}$	$4.17X10^{02}$
CD27	-1.55	$9.78X10^{05}$	$4.37X10^{02}$

B.



C.

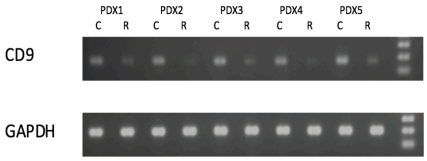


Figure 6.6. Differentially expressed genes in relapsed PDXs

(A) Differentially expressed genes with FC >1.5 and FDR <0.05 were identified and ranked by FC. (B) The top eight differentially expressed genes in both categories were validated by RT-qPCR across all 5 PDXs and representative graphs are shown for ABCB1 and STXBP6 with means and standard deviations of technical triplicates. (C.) RT-PCR was carried out on a select set of genes and a representative figure for CD9 highlights the decreased transcript level in each PDX after treatment, C = control, R = relapse.

A.

Gene Set Name [# Genes (K)]	# Genes in Overlap (k)	k/K	p-value	FDR q-value
GO_POSITIVE_REGULATION_OF_MITOTIC_CELL ELL_CYCLE	3		$5.25\ e^{-6}$	$8.75\ e^{-3}$
GO_RESPONSE_TO_ABIOTIC_STIMULUS	5		$6.03\ e^{-6}$	$8.75\ e^{-3}$
GO_CELL_CYCLE_PROCESS	5		$7.85\ e^{-6}$	$8.75\ e^{-3}$
GO_ORGANELLE_FISSION	4		$8.87\ e^{-6}$	$8.75\ e^{-3}$
GO_CELL_CYCLE	5		$2.03\ e^{-5}$	$1.48\ e^{-2}$
GO_CELLULAR_RESPONSE_TO_EPIDERMAL_GROW ROWTH_FACTOR_STIMULUS	2		$2.21\ e^{-5}$	$1.48\ e^{-2}$
GO_REGULATION_OF_CELL_ADHESION	4		$2.25\ e^{-5}$	$1.48\ e^{-2}$
GO_RESPONSE_TO_EPIDERMAL_GROWTH_FACTOR	2		$3.2\ e^{-5}$	$1.83\ e^{-2}$
GO_POSITIVE_REGULATION_OF_PROTEIN_META ETABOLIC_PROCESS	5		$3.71\ e^{-5}$	$1.83\ e^{-2}$
GO_REGULATION_OF_CELL_DIFFERENTIATION	5		$3.71\ e^{-5}$	$1.83\ e^{-2}$

B.

Gene Set Name [# Genes (K)]	# Genes in Overlap (k)	k/K	p-value	FDR q-value
GO_EXTRINSIC_APOPTOTIC_SIGNALING_PATHW THWAY	3		$1.15\ e^{-6}$	$6.81\ e^{-3}$
GO_POSITIVE_REGULATION_OF_ESTABLISHMEN MENT_OF_PROTEIN_LOCALIZATION	4		$3.08\ e^{-6}$	$7.27\ e^{-3}$
GO_REGULATION_OF_CELLULAR_LOCALIZATION	5		$3.69\ e^{-6}$	$7.27\ e^{-3}$
GO_ACTIN_FILAMENT_ORGANIZATION	3		$6.28\ e^{-6}$	$8.51\ e^{-3}$
GO_REGULATION_OF_CELL_DEATH	5		$7.38\ e^{-6}$	$8.51\ e^{-3}$
GO_PROTEIN_COMPLEX_SUBUNIT_ORGANIZATIO TION	5		$8.82\ e^{-6}$	$8.51\ e^{-3}$

Figure 6.7. Gene ontology analysis of differentially expressed genes following relapse

Significantly (A) up- and (B) down-regulated genes were compared to GO gene sets in MSigDB and the top enriched biological processes for each are listed (FDR < 0.02).

6.2.6 A subset of PDXs show upregulation of B cell receptor and PI3K signalling after chemotherapy treatment

In order to investigate inter-tumour heterogeneity amongst the five untreated control patient samples, principal component analysis (PCA) was carried out. These data highlight the clustering of three PDXs more closely to one another, with the remaining two PDXs separated both from each other and from the three clustered samples (Figure 6.8). Based on the greater similarity of PDXs 2, 3 and 5 to one another, differential gene expression analysis was repeated comparing just these three control PDXs to their matched relapsed samples. This highlighted a greater number of differentially expressed genes with 384 genes with a *p* value <0.05 (Figure 6.9A). Gene set enrichment analysis (GSEA)[374] was carried out using normalized mRNA read counts for the six samples with gene sets in MsigDB[533], particularly Hallmark and KEGG gene sets. Analysis of enriched gene sets highlighted that both the PI3K and B cell receptor signalling pathways were enriched in relapsed samples with a *p* value = 0 and a FDR <0.2 (Figure 6.9B).

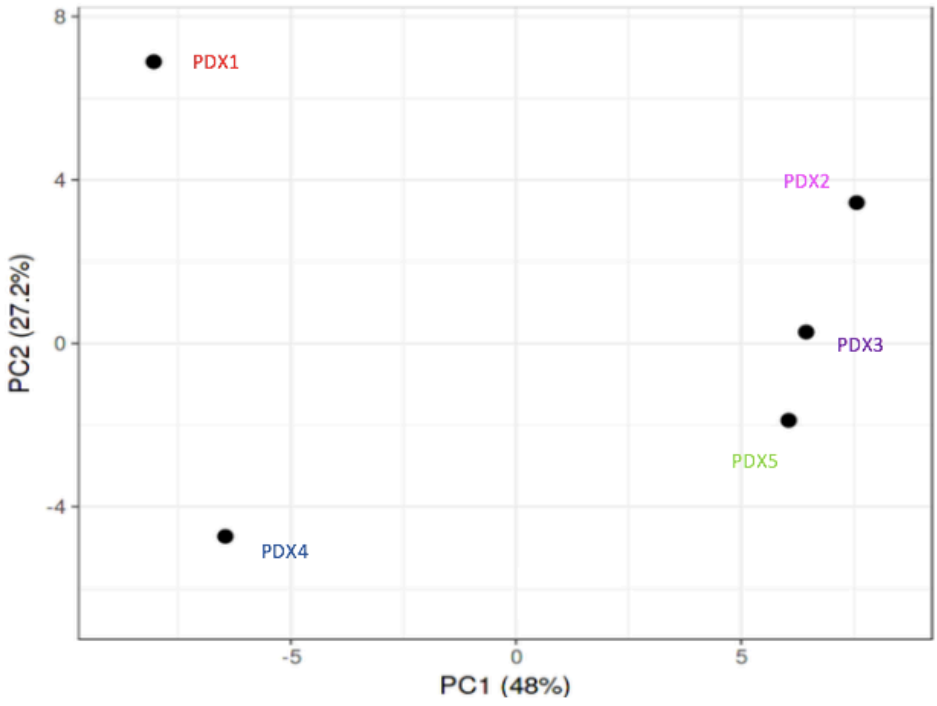


Figure 6.8. Principal component analysis demonstrates inter-tumour heterogeneity

Principal component analysis was carried out based on gene expression profiles of each PDX. X and Y axes show principal component 1 and principal component 2 that explain 48% and 27.2% of the total variance, respectively.

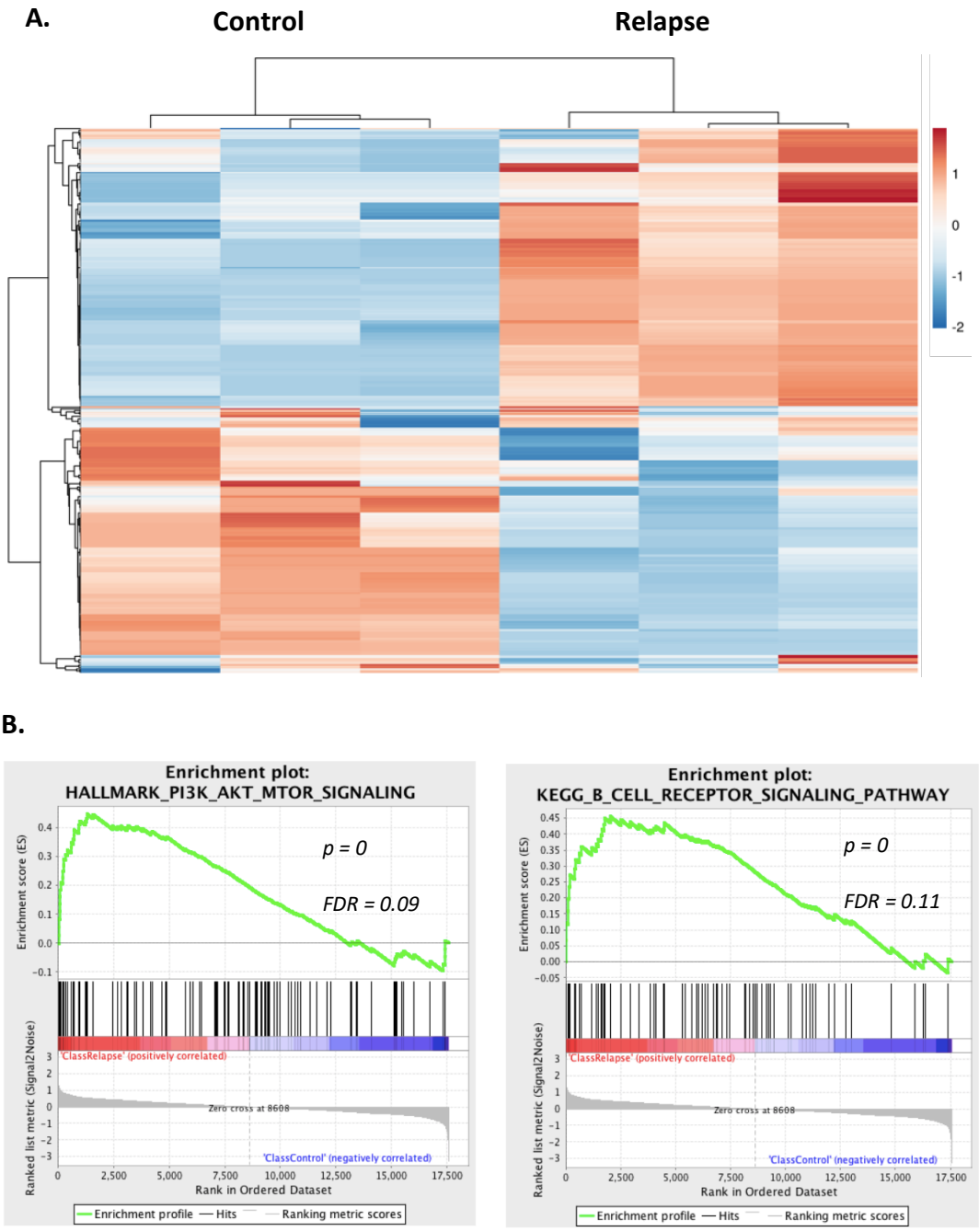


Figure 6.9. Differential gene expression analysis of relapse and control PDX 2, 3 and 5

(A) Hierarchical clustering analysis was carried out using normalized read counts of differentially expressed genes (p <0.05) for each sample. Red represents high, and blue low relative expression. (B) GSEA was performed to rank the genes based on their differential expression levels between control tumours and relapsed tumours. A deviation from an enrichment score (ES) of 0 in the profile (green line) reflects the degree of over-representation of genes at the top (positive ES score) or bottom (negative ES score) of the ranked gene list.

6.3 Discussion

Given the dismal prognosis of relapse and progressive disease in paediatric B-NHL, efforts to gain an understanding of the mechanisms of resistance are paramount. The results presented in this chapter lead to a number of interesting observations with respect to the response of PDXs to chemotherapy. For example, PDX 1 and 4 were derived from patients similarly presenting with stage III disease involving pleural fluid from which the PDXs were generated. Despite these similarities, the PDX showed differential response to *in vivo* therapy (PDX4 was less chemo-responsive requiring increased drug doses compared to PDX1) supportive of inter-patient variability despite comparable initial disease presentation. The heterogeneity amongst the small PDX sample group in this study is reflective of a large scale study of 1000 PDXs developed from solid tumours, whereby inter-patient heterogeneity in response to therapies within the same cancer type was observed[534]. Interestingly, this study and others have been able to show that PDX models respond to therapy similarly to the patients from which they are derived, highlighting the faithfulness of the PDX model and the potential significance of the observed clinical responses[399,534–537]. Unfortunately, it has not been possible to compare patient clinical response to therapy, to the derived PDX outcome in this thesis, as clinical data were unavailable.

Data presented in chapter 4 shows that SP cells isolated from each of the five PDXs have increased tumourigenic potential compared to MP cells, as well as increased expression of ABC transporters. Apart from common upregulation of ABC transporters the SP fraction isolated from each PDX has a distinct gene expression pattern. Genes associated with anti-apoptotic programmes were upregulated across the different PDXs and to determine if the detected gene expression changes specific to the SP (as compared to the MP) are functionally relevant to treatment resistance, whole cell and SP-derived tumours of PDX1 were treated with chemotherapy *in vivo.* The relative, diminished response of the four SP-derived tumours to chemotherapy compared to whole population-derived tumours is significant, highlighting the advantage SP cells have in resisting chemotherapy-induced cell death. The identification, isolation and functional verification of a heterogenous population within a PDX that has increased resistance to chemotherapy is an important insight into pBL. Increased expression of ABC

transporters in the SP cells is an obvious likely contributor to chemoresistance, although the RT-qPCR analysis presented in chapter 4 identifies other differentially expressed genes in the SP of PDX1 which may also be responsible for this effect, including AKT, MCL1 and PIM1 (Figure 4.12). Each of these have been identified as mediators of stemness and therapy resistance in other studies and further investigation of their contribution to maintaining the SP phenotype within the PDXs presented in this thesis would be informative[466,482,538–540].

Due to both the rarity of relapse and the lack of available relapse biopsies, it is difficult to study the biology underlying relapse in pBL. The PDXs developed in this thesis therefore provide a model system in which to 'mimic' relapse *in vivo*. The validity of the PDX models in achieving this aim is evidenced by the induction of a period of 'remission' after intense chemotherapy followed by re-growth, somewhat mimicking a clinical scenario. However, there are some limitations to this model which should be considered when interpreting the data produced. For example, a 'relapsed' tumour was produced for all five PDXs, whereas clinically, less than 10% of pBL patients experience relapse suggesting that the chemotherapy level in the PDX model was sub-optimal. However, chemotherapy had a significant impact on the welfare of the mice and it was necessary to ensure that clinical signs did not exceed acceptable severity limits such as a weight reduction of more than 20%. As such, it was not possible to continue chemotherapy during equivalent 'remission' as is done in the consolidation stage of human chemotherapy regimens. Hence, to account for the 100% rate of relapse observed in the PDX model, it is likely the high dose chemotherapy received clinically by patients is sufficient to eliminate the SP fraction identified in our PDX model if it also exists in all pBL cases, i.e. it is only in rare BL cases that this population is able to prevail and result in treatment failure. Indeed, in favour of this theory there are distinct molecular differences between the SP of the individual PDXs. If the SP is the cause of resistance in true clinical cases, it would be imperative to determine if this population have unique characteristics that make them more resistant than the SP in treatment-responsive patients. Even though our model does not fully recapitulate the intensity of treatment received by patients, it does confer a strong selective pressure on tumour cells which ultimately gives rise to a dramatic reduction in tumour size and its subsequent re-

growth. As such, the increase in SP frequency in each of the relapse PDXs, despite still constituting a minority of the tumour, is an interesting observation. The most obvious reason for this increase is that SP cells, with their increased resistance to chemotherapy, are preferentially selected during chemotherapy. In time they give rise to a new tumour consisting of SP and MP cells, re-establishing an equilibrium as observed previously in chapters 4 and 5 for both PDX and cell line SP sorted cells. However, whether there is a gradual and relative decrease in SP cells as the tumour regrows and if indeed the remaining cells following chemotherapy are exclusively SP remains to be determined. Alternatively, residual cells following initial chemotherapy may be MP cells that reacquire stemness characteristics through either genetic or epigenetic aberrations (more likely the latter given the short time frame) giving rise to SP and ultimately MP subsets. In support, Gross *et al.* have previously shown in CLL that chemoresistance involves 'innate' and 'acquired' leukemic SP cells[541]; SP cells showed preferential survival compared to non-SP cells following treatment with Fludarabine, Bendamustin and Rituximab. They also highlighted that it wasn't exclusively the SP cells that drove resistance but that drug treatment also resulted in acquired SP cells and although the SP still made up the minority of tumour cells, after one year of treatment patients had between a three and 100 fold increase in the SP fraction[541]. A recent publication by Schmitt and colleagues also highlighted mechanistic insights into the plasticity of the CSC in lymphoma relapse[542]. Significantly, temporary enforcement of senescence in p53-regulatable models of ALL and AML was shown to reprogram non-stem bulk leukaemia cells into self-renewing, leukaemia-initiating stem cells. This was determined by increased expression of Kit and Sca1 leukaemic stem cell markers as well as transcript levels of stemness genes including Abcg2 and Cd44. They suggest that Wnt signalling is an essential driver of the stemness exhibited by tumour cells after release from senescence and in patient matched presentation and relapse samples of DLBCL they identified activated Wnt signalling after relapse. This unexpected link between therapy induced senescence and stemness has important therapeutic implications and further highlights the complexity of the stemness state. Given evidence in this chapter, that the SP plays an important role in chemo-resistance, it will be essential to further elucidate the exact dynamics involved. Further experiments might include lineage-tracing by

'barcoding' cells to map the clonal dynamics of SP cells compared to the bulk of the tumour upon treatment with chemotherapy.

Even though barcoding and lineage tracing was not conducted in this study, differential expression profiles between the whole untreated and treated PDXs was conducted in a pilot RNA-seq study of the five matched PDX pairs (primary PDX tumour and relapse) with differentially expressed genes identified and successfully validated across all samples. GO analysis of the limited set of differentially expressed genes highlighted that they were largely associated with cell cycle progression and apoptotic pathways suggesting that evasion of apoptosis is an important response to chemotherapy in BL. CD9 was also one of the most differentially expressed genes across all 5 PDXs. CD9 negativity has been noted as significant throughout this thesis with CD9 negative cells identified as having increased tumourigenic potential in chapter 4 and CD9 surface expression was shown to be lower in the SP fractions of PDXs. Determining the relevance of CD9 expression in the context of tumour progression is difficult as its underlying functions are not well known and opposing roles have been reported in the literature. It is a membrane glycoprotein that belongs to the tetraspanin family characterized by four transmembrane domains[543]. Through partnering with integrins and immunoglobulin superfamily receptors on the cell surface it transduces extracellular signals inside to promote apoptosis, suppress metastasis, and mediate cell adhesion in a number of studies[464,544–547]. Downregulation of CD9 has also been associated with metastasis in pancreatic and ovarian cancer[30,260]. Yoon *et al.,* in a study of B-cell lymphomas showed that CD9 positive cells were more susceptible to apoptosis than CD9 negative cells and that its expression was epigenetically regulated[464]. They also highlighted that in the transformation from indolent FL to DLBCL, patients who had high initial CD9 expression had lower expression following transformation to the malignant form. However some cases with low original CD9 expression showed no change[464]. A recent study in FL also showed that CD9 expression was associated with better outcomes and that CD9 expression negatively correlated with PIK3CD expression i.e. $CD9^-/PIK3CD^+$ patients had the shortest PFS, suggesting a potential link between CD9 and PI3K signalling[543]. However, contrary to these reports, a recent study in paediatric B-ALL suggested that CD9 expression is associated with adverse survival outcomes[548].

Given the recurring significance of this marker throughout this project and the mentioned literature, it is likely that CD9 has an important role in BL progression and should be further studied.

Another gene upregulated in all five PDX relapse samples was MALAT1, which is a particularly interesting observation based on the emerging role of long non-coding RNAs (lnRNAs) in lymphomagenesis. Furthermore, MALAT1 is associated with tumour metastasis, progression, drug resistance and stemness in a host of malignancies with evidence supporting its role in cell cycle progression as well as Wnt and MAPK signalling pathways[549–552]. Indeed, its increased expression is associated with a worse OS in NSCLC, renal cell carcinoma, colorectal cancer, MCL and T cell lymphoma and it has been reported to be upregulated in response to methotrexate[553–558]. Interestingly, in independent studies, MALAT1 has also been shown to maintain the undifferentiated state of early-stage haematopoietic cells and to contribute to the maintenance of stemness in breast CSCs by regulating Sox2 and CD133[559–561]. Moreover, one explanation for the reduced overall survival associated with MALAT1 in MCL is due to its ability to increase expression of EZH2, a member of the Polycomb Repressive Complex 2 (PRC2), resulting in cell proliferation[554]. Whether MALAT1 plays a similar role in BL remains to be investigated but EZH2 has been shown to be deregulated in BL by both activating mutations and repression of its negative regulator by MYC[52,562].

Due to the paucity of differentially expressed genes identified and the potential for inter-tumour heterogeneity to mask functionally significant results, PCA was conducted. These data highlight a degree of separation between PDXs with regard to their gene expression profiles whereby three PDX samples showed relative similarity, clustering separately from the other two models. As such, gene expression data from the three more closely related PDXs were analysed separately identifying upregulation of both the PI3K and B cell receptor signalling pathways in relapse tumours. The identification of increased PI3K signalling in these cases, after chemotherapy treatment is particularly interesting in light of the identified role for PI3K/mTOR signalling in maintaining the SP phenotype in BL cell lines in chapter 5. It is possible that the increased expression of genes in these pathways is due directly to the higher SP fraction in the relapse PDXs,

although it is not possible to determine the relative contribution of individual cells to the overall gene expression profile, but as they constitute approximately 5% of the bulk tumour, this is unlikely to be the case. It is more likely that MP cells which may (or may not; see discussion above) have been produced from surviving SP cells after chemotherapy display a PI3K signalling gene expression profile as a consequence of their recent differentiation from SP cells, maintaining elements of SP signalling or alternatively, sustained exposure to chemotherapeutic agents results in up-regulation of this pathway regardless of an SP or MP derivation.

Increased expression of B cell receptor signalling on relapse of the PDX tumours is also an interesting finding, particularly as inhibitors of this pathway such as Ibrutinib have been successful in the clinic in the treatment of relapsed or refractory adult B-cell malignancies[108]. A clinical trial investigating Ibrutinib in relapsed paediatric B-NHL is currently underway and it will be interesting to see the initial results in the coming months. Of note, B cell receptor signalling was not specifically investigated in our SP cells, but treatment of cell lines in chapter 5 with Ibrutinib did not result in a significant decrease in SP cells. This highlights that the enrichment of this pathway may not be due to the increased SP fraction but rather may be upregulated in cells, including SP cells, after chemotherapy.

Taken together this chapter offers a number of different insights to both intra- and inter-tumour heterogeneity. We have demonstrated that one of the heterogenous populations identified in this thesis, the SP, has significantly increased resistance to chemotherapy and appears to play an important role in relapse based on its increased frequency in tumours after treatment. The ability to utilise PDXs to mimic patient treatment is also a useful resource which highlights the possibility that individual tumours may respond differently to chemotherapy. Indeed, despite some common differentially expressed genes involved in drug efflux and cell cycle progression, three PDXs showed upregulation of PI3K and B cell receptor signalling that was not present in the other two relapsed PDXs. This study is not based on clinically true relapse samples and includes a small sample number although it does offer preliminary data suggesting response to chemotherapy in patients may not be uniform. A recent landmark paper

from the Staudt group has identified new subgroups of DLBCL with distinct genotypic, epigenetic, and clinical characteristics, providing potential for precision-medicine strategies in DLBCL[563]. A similar determination in pBL would be difficult, particularly due to the limited sample size, although this type of study is encouraging in the B-NHL field and highlights that it is possible to further categorise patients to predict their response to chemotherapy.

Chapter 7. Clustered regulatory interspaced short palindromic repeats (CRISPR)-Cas9 overexpression screen to identify putative resistance mechanisms to chemotherapy and Rituximab

7.1 Introduction

In previous chapters the high clinical relevance of our PDX models has been highlighted, and in particular they are an excellent resource for investigating potential therapeutics for BL. However, PDXs are not an infinite resource, particularly as efforts should be made to utilise low passage numbers for experiments. For this reason modern genomic and functional analytical techniques should first be utilised to determine novel genes and signalling networks that are involved in mediating resistance of tumour cells to a particular treatment. Identified rational targets can then be investigated in the PDX models. In this chapter a genome-wide CRISPR-Cas9 (CRISPR-associated protein 9) overexpression screen was performed to identify putative resistance mechanisms to therapeutic agents currently used in the frontline treatment of pB-NHL. FAB/LMB protocols incorporate numerous therapeutic agents - the group B backbone chemotherapy consists of cyclophosphamide, vincristine and prednisone followed by two induction cycles which incorporates these drugs along with methotrexate and finally consolidation cycles consisting of cytarabine plus methotrexate. Group C high-risk protocols include each of these drugs along with etoposide for two cycles of intensification[21]. The BFM backbone is similar to the FAB/LMB regimen but also incorporates ifosfamide. Due to the large number of agents used it was not feasible to investigate all of them in our studies therefore three different agents, each with a different mechanism of action, were chosen for this screen; (1) doxorubicin, a topoisomerase II inhibitor, (2) methotrexate, an antimetabolite and (3) vincristine, an inhibitor of microtubulin assembly. Despite the efficacy of these chemotherapeutic regimens resistance still poses a significant challenge. Examples of resistance mechanisms can include increased drug efflux, mutations of the drug target, DNA damage repair, activation of alternative signalling pathways and evasion of cell death and were discussed in chapter 1 with respect to these drugs[512]. A genome-wide screen offers the opportunity to identify novel resistance mechanisms that may be therapeutically relevant. The introduction of Rituximab to front line therapy has been

highly effective although relapse cases still occur and for this reason Rituximab was also investigated in this screen[102]. The main mechanisms of action of Rituximab include ADCC and CMC while direct induction of apoptosis has also been widely reported[564]. We acknowledge that the *in vitro* model of resistance in this chapter does not address ADCC or CMC mechanisms of resistance and any identified mechanisms of resistance to Rituximab would require validation in further systems, for example in our PDXs models established in humanised mice.

Since their introduction, genome-wide forward screens have proven powerful tools for the unbiased discovery and functional characterization of specific genetic elements associated with a phenotype of interest[565]. Given the relative ease of generating libraries containing thousands of single guide RNAs (sgRNAs) that can be used to create large mutant collections, CRISPR-Cas9 technology has quickly become the method of choice for large-scale pooled screens. In particular, a pooled screening approach provides an opportunity to interrogate thousands of genetic perturbations in a single experiment. Initially, the majority of pooled screens performed utilised wild-type (WT) Cas9 to perform CRISPR-Knock Out (KO) screens. However, an increasing number of studies now also utilize the catalytically-dead mutant of Cas9, referred to as dead-Cas9 or dCas9, in which the nuclease activity of the WT Cas9 has been mutated in order to render it non-functional. dCas9 has been fused to a range of chromatin modifier fusion proteins to convert it into a highly versatile enzyme that can be used to perform activation (CRISPR activation (CRISPRa)) or repression (CRISPR interference (CRISPRi)) screens[566]. Critics of the screen approach will label it a 'fishing expedition' and indeed, like most omics approaches, genome-wide screens do not fit the hypothesis-driven research model. However, numerous studies have highlighted the power of this method. The Weissman group adapted dCas9-based activators and repressors to carry out powerful genome-wide gene activation and repression screens which showed accurate complementarity[567.] A subsequent study by Feng Zhang's lab also demonstrated the successful adaptation of dCas9-based activators for genome-wide gene activation screens and notably also demonstrated the feasibility of identifying mediators of resistance to a BRAF inhibitor[375]. In positive selection screens, a strong selective pressure is introduced such that there is only a low probability that cells without a

relevant survival-enhancing perturbation will remain following selection. Usually positive selection screens are devised to identify perturbations that confer resistance to a drug, toxin or pathogen.

CRISPRa in particular offers a significant improvement as a screening platform over other activation approaches. Previously gain of function (GOF) screens were primarily limited to cDNA overexpression libraries, which suffered from incomplete representation, difficulty cloning or expressing large sites of DNA construct, lack of isoform diversity, and high cost of construction[568]. CRISPRa overcomes these limitations because it activates gene transcription at the endogenous locus and simply requires the synthesis and cloning of specific RNA guides, making it much more affordable.

This project is based on the Synergistic Activation Mediator (SAM) overexpression CRISPR library (Table 7.1), to identify GOF events that may induce resistance to both a standard chemotherapy backbone and/or Rituximab.

	SAM
Induced phenotype	**Overexpression**
Number of genes targeted	**23,430 (all coding isoforms)**
sgRNAs per gene	**3**
Total sgRNA constructs	**70,290**

Table 7.1. Key characteristics of the SAM screen library

The SAM library is based on a dCas9-VP64 fusion protein that recruits transcriptional complexes to the transcription start site (TSS) of target transcripts. SAM improved upon the first generation of CRISPR GOF strategy by altering sgRNA to also recruit transcriptional co-activators to synergistically interact with the transcriptional complex recruited by dCas9-VP64[375,569]. Specifically, a hairpin aptamer that selectively binds to MS2 phage protein was appended to the tetraloop and stem loop #2 regions of the sgRNA, which were identified to be exposed outside of the Cas9 enzyme based on the crystal structure of the Cas9-sgRNA complex[377]. Then, a separate vector was constructed to express MS2 protein fused to p65 transcription factor and heat shock transcription

factor 1 (HSF1). This allowed for additional recruitment of transcriptional activators to the TSS, leading to significantly (sometimes >500-fold increase in targeted mRNA level vs. dCas9-VP64 alone) enhanced overexpression capability for genes that were considered to be difficult to overexpress with the previous CRISPR overexpression (OV) systems[375]. For this reason, the SAM screen library was selected for this investigation. This chapter overviews preparation for the screen as well as quality control and validation of the data followed by analysis of clinical relevance of the identified candidate genes. A diagram of the screen process is shown in Figure 7.1.

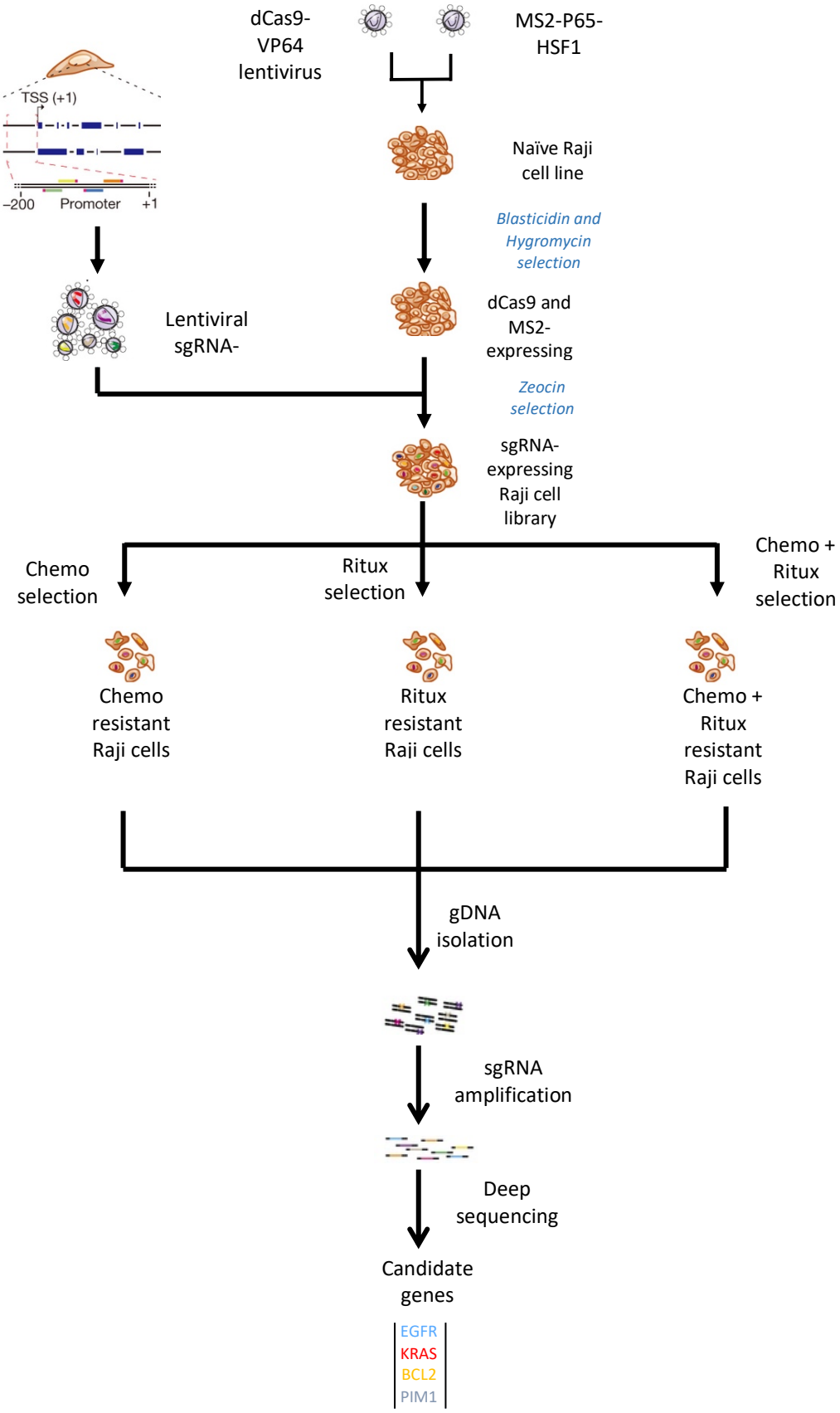

dCas9-
VP64
lentivirus
MS2-P65-
HSF1
TSS (+1)
−200
Promoter
+1
Naïve Raji
cell line
Blasticidin and
Hygromycin
selection
Lentiviral
sgRNA-
dCas9 and
MS2-
expressing
Zeocin
selection
sgRNA-
expressing
Raji cell
library
Chemo
selection
Ritux
selection
Chemo +
Ritux
selection
Chemo
resistant
Raji cells
Ritux
resistant
Raji cells
Chemo +
Ritux
resistant
Raji cells
gDNA
isolation
sgRNA
amplification
Deep
sequencing
Candidate
genes
EGFR
KRAS
BCL2
PIM1

Figure 7.1. Key characteristics of the SAM screen library

Genome-scale Cas9 transcriptional activation screens begin with the construction of a plasmid library encoding the effector protein and sgRNAs. The SAM genome-scale sgRNA library consists of 70,290 guides, targeting every coding isoform from the RefSeq database (23,430 isoforms, three guides per isoform). The plasmid library is packaged into lentivirus and then transduced into the cell type of interest to generate stably expressing lines for the screen, along with an accessory transcriptional activator complex. A selection pressure – e.g. Rituximab – is applied and genomic DNA is harvested. Surviving sgRNA sequences (coloured bars) are amplified from genomic DNA and then analysed by deep sequencing to identify candidate genes. Individual candidate genes are then validated. Chemo = a standard chemotherapy backbone, Ritux = Rituximab.

Aims of this chapter

- Determine concentrations of drugs to use for multiagent treatment in the screen
- Perform quality control analysis of the screen results (NGS readout) at the sequence, read count and sample levels
- Perform multiple data analysis to obtain a list of candidates that may induce resistance to chemotherapy and/or Rituximab
- Perform individual validation (viability-based) assays for candidate genes to confirm their capability to induce resistance to chemotherapy and/or Rituximab
- Identify the clinical relevance of candidate genes.

7.2 Results

7.2.1 Validation of CRISPR/dCas9-VP64 induced overexpression phenotype in the Raji cell line

Before validating SAM functionality, the successful transduction of dCas9 into the Raji cell line was verified by Western blot (Figure 7.2). Following this, the functionality of the SAM library three-vector overexpression system was investigated by targeting 15 genes (2 sgRNAs per gene) in the SAM library-ready (stably expressing dCas9-VP64 and MS2-P65-HSF1) Raji cell line. These genes encode for growth factors, guanine nucleotide exchange factors, transcription factors, kinases and adapter proteins. The SAM system was validated to be functional in Raji cells as the target genes were overexpressed at a fold-difference range of 2 to 3200 at the mRNA level (Figure 7.3). The 15 genes included in the assay are ones that are expressed at low, medium and high endogenous levels in B-NHL cells and so allowed for the relationship between endogenous expression and SAM-induced expression to be analysed.

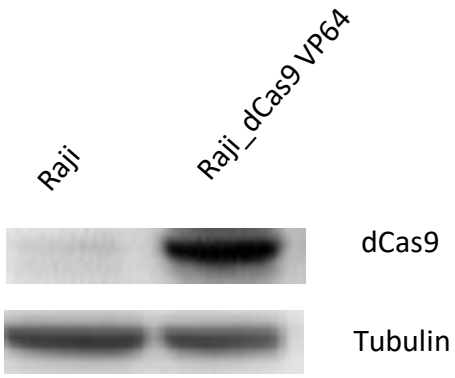


Figure 7.2. Successful transduction of dCas9 in Raji cells was verified by Western blot

Following transduction of dCAS-VP64_Blast and lenti MS2-P65-HSF1_Hygro into Raji cells, protein lysates from transduced and untransduced control Raji cells were prepared and Western blot carried out to detect dCas9 and tubulin.

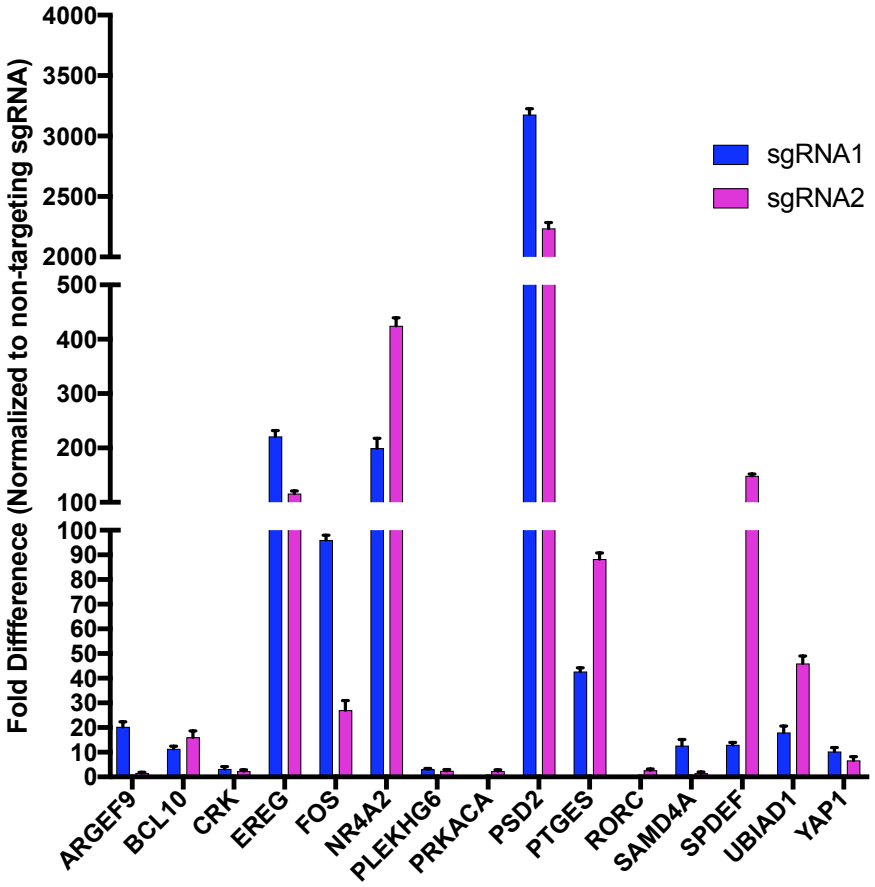


Figure 7.3. The SAM system induces overexpression of various endogenous genes in Raji cells.

RT-qPCR was conducted five days after sgRNA transduction (2 sgRNAs per gene) in technical triplicates and ΔΔCT analysis was conducted with GAPDH as an internal control with normalization to the non-targeting (NT) sgRNA sample. Means and standard deviations of technical triplicates are shown.

Although the functionality of the SAM system was confirmed, it is clear from these data that overexpression (OV) of some genes is to a higher level than others; six genes (FOS, *SPDEF, PSD2, PTGES, EREG, NR4A2*) were overexpressed at a much higher level compared to the rest of the target genes (Figure 7.3). Furthermore, four genes (*PLEKHG6, CRK, RORC, PRKACA*) showed very little response to the SAM system's OV induction (Figure 7.3). To determine whether this was due to certain genes already being highly expressed, the endogenous expression levels of each of these genes in the Raji

cell line was obtained from the Cancer Cell Line Encyclopaedia[570], and compared to the fold change induction but these data did not correlate (Figure 7.4), suggesting the variation in the SAM system's OV induction is due to inherent characteristics of this library, most likely specificity and efficiency of the sgRNAs.

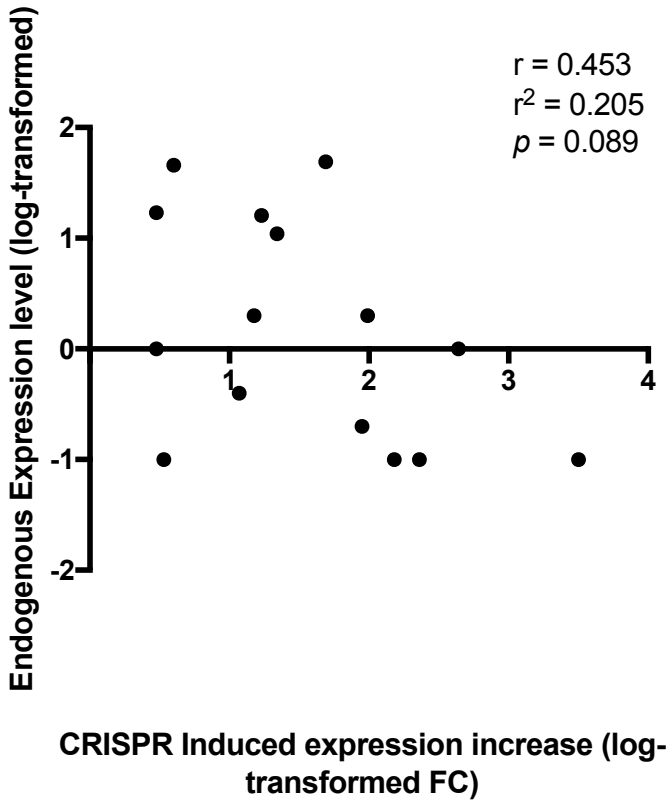


Figure 7.4. Pearson correlation between the endogenous unmodified expression level of the target genes and the level of SAM-mediated overexpression induction.

Both the endogenous expression levels and the fold change in gene expression induced by SAM were log-transformed, p = 0.089.

7.2.2 Determining optimal drug concentrations for individual arms of the screen

In order to determine the optimal drug concentrations to use in the screen, a bliss independence model was used to determine which concentrations of drugs gave synergistic, additive or antagonistic effects. The Bliss Combination Index (BCI) can be calculated with the following equation: $BCI = \frac{Ea+Eb-Ea*Eb}{Eab}$, where Ea indicates the viability effect of drug A (in this case, doxorubicin) and Eb indicates the viability effect of drug B (in this case, methotrexate) and Eab indicates the viability effect of being treated with drug A and drug B together[365]. As such, a BCI less than 1 indicates synergism, a BCI equal to 1 suggests an additive effect and a BCI greater than 1 is indicative of antagonism between two drugs. Initially the dose response curve for each drug was carefully determined using a range of eight concentrations of drug and non-linear regression methods (Figure 7.5). Once the ED50 had been determined, the ED20, ED25, ED30 and ED35 for each drug was calculated. Employing a matrix approach, combinations of the drugs were then applied to the cells to generate a heat map of BCI values (Figure 7.6A). As indicated, there is no significant synergy nor antagonism associated with any combination of drugs with most BCI values being close to one indicating additive effects. Nonetheless, the two combinations that gave the lowest BCI values were selected (doxorubicin ED30 + methotrexate ED25 and doxorubicin ED30 + methotrexate ED30) and used as fixed doses for assessing the effects of combining with four different doses of vincristine (Figure 7.6B). Given the observed synergy, albeit at a low level, the doxorubicin ED30, methotrexate ED30 and vincristine ED35 dosages were chosen for the chemotherapy arm of the screen. Cell viability assay after treatment of Raji cells with these concentrations of drugs determined an ED of 70. The theoretical value using the enumerator of the BCI formula would predict an ED of 68, highlighting that there is some minor synergy observed (BCI = 0.97). In order to determine the effective dose for Rituximab, the combination of doxorubicin at the ED30, methotrexate at the ED30 and vincristine at the ED35 was applied to the cells together with Rituximab at a range of doses from the ED20 to the ED40 (Figure 7.6C). There was a minor synergistic effect for the combination of Rituximab at its ED40 with the fixed chemotherapy dose. Therefore, the dosage used for the chemotherapy + Rituximab arm of the screen was doxorubicin ED30, methotrexate ED30, vincristine ED35 and

Rituximab ED40. Following treatment of Raji cells at these dosages, an ED value of 85 was observed, slightly greater than the mathematically predicted ED value of 80 giving a BCI value of 0.94.

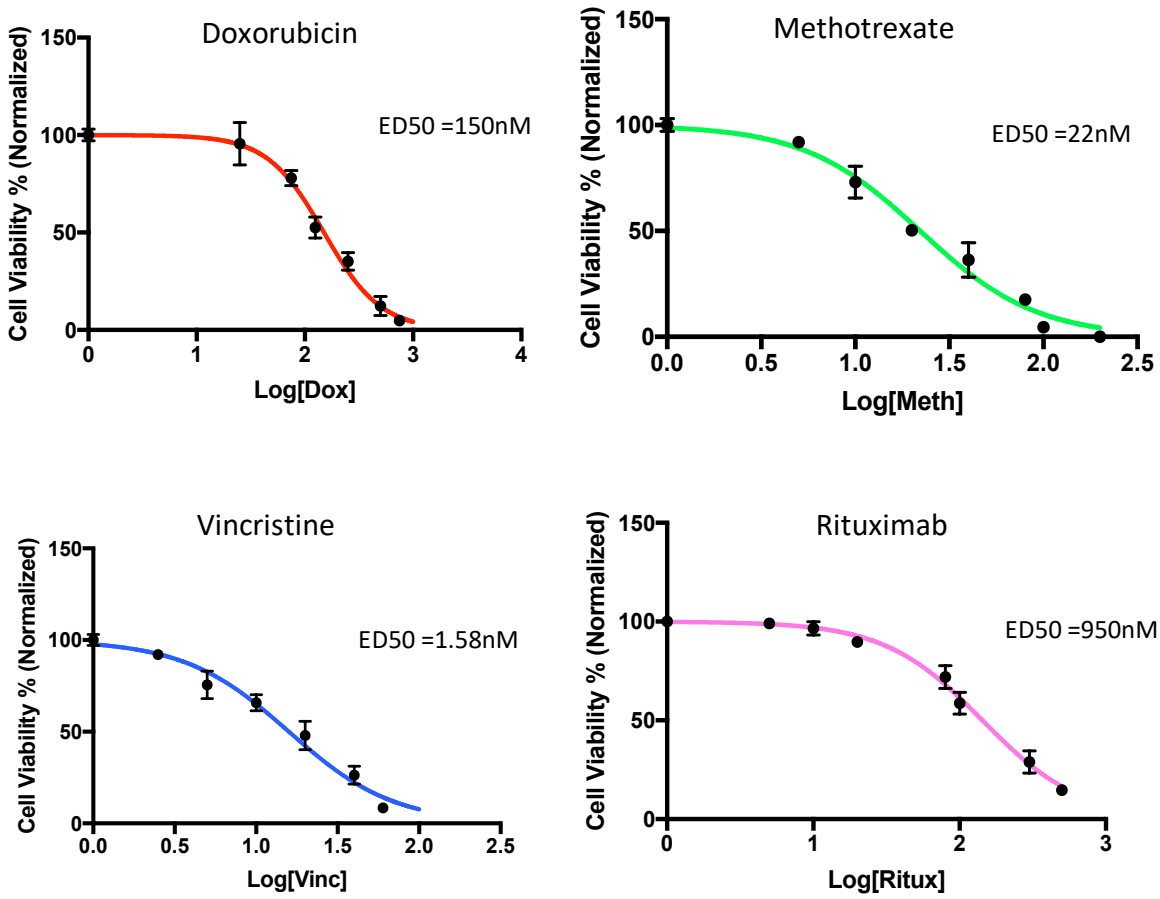


Figure 7.5. Dose response curves for individual drug treatments

Raji cells were treated with increasing concentrations of doxorubicin, methotrexate, vincristine or Rituximab for 72 hours. End point viability compared to untreated control was determined by the CellTiter-Blue viability assay. ED50s are listed beside the non-linear regression plot displaying mean viabilities and the 95% CI of technical triplicates.

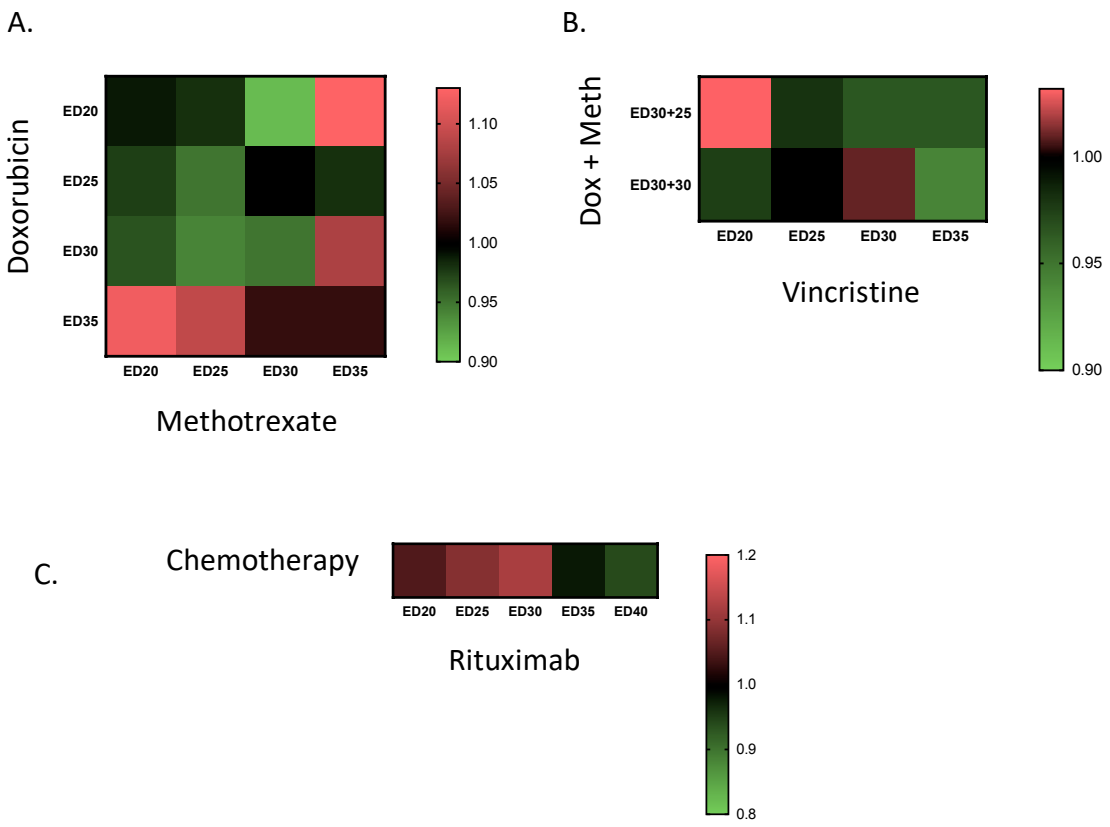


Figure 7.6. Bliss independence analysis to determine possible synergistic drug combinations.

Raji cells were treated in a matrix format with doxorubicin and methotrexate at concentrations that equated to the ED20, ED25, ED30 and ED35 for each drug for 72 hours in triplicates (A), doxorubicin at the ED30 and methotrexate at the ED25 as well as doxorubicin at the ED30 and methotrexate at the ED30 were applied to Raji cells together with vincristine at the ED20, ED25, ED30 or ED35 (B) or doxorubicin at the ED30, methotrexate at the ED30 and vincristine at the ED35 were combined with Rituximab at its ED20, ED25, ED30, ED35 and ED40 (C). End point viability compared to the untreated control was determined by CellTiter-Blue viability assay and the BCI determined for each drug combination and displayed in a heatmap.

7.2.3 Quality control analysis of the SAM screen

The SAM screen was performed as illustrated in Figure 7.1. Briefly, the sgRNA library was transduced into ~2 billion Raji cells at ~0.3 MOI and the transduced cells were then cultured in the presence of DMSO (negative control) or drugs (chemotherapy, Rituximab or chemotherapy + Rituximab) at the chosen ED values (chemotherapy = ED70, Rituximab = ED70, chemotherapy + Rituximab = ED85) for 12 days before genomic DNA was extracted from all samples and the sgRNA region amplified by PCR to generate Illumina-ready libraries for analysis on a HiSeq platform (conducted at Bauer Core, Harvard University).

Before conducting any gene-based analysis to identify putative resistance mechanisms, three layers of quality control (QC) analysis were performed with MaGECK-VISPR to ensure that the screens were adequately executed. Specifically, QC analysis of all raw fastq files derived from the screen samples was conducted at the sequence, read count and sample levels.

The sequence level QC was designed to identify potential technical issues associated with the sequencing process itself and was conducted with FastQC script (Babraham Bioinformatics) that was embedded in MaGECK-VISPR. Sequence level analysis of SAM screen samples satisfied the general QC guidelines (Figure 7.7).

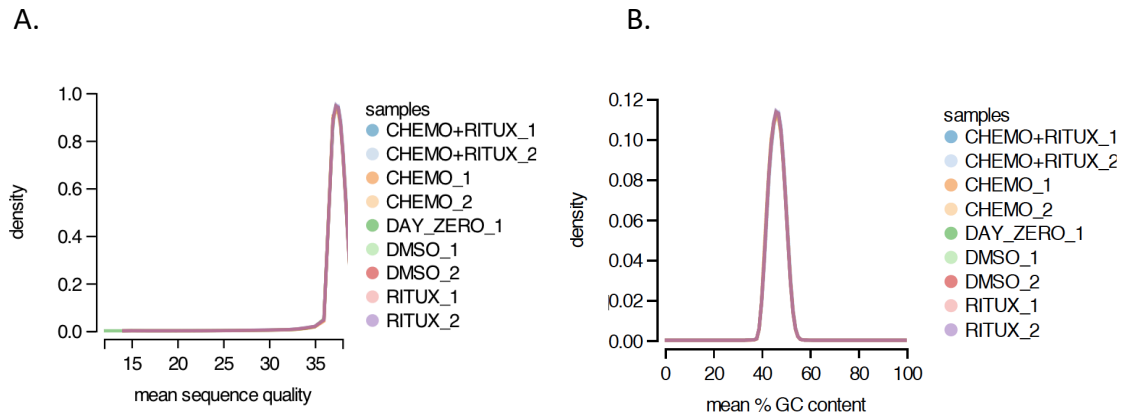


Figure 7.7. Sequence level QC of SAM screen samples.

(A) Distribution of per-read 'mean base quality' for each sample. All samples from the SAM screen overlap perfectly with a peak at ~38, which is well beyond the general threshold value of 30. (B) Per sample distribution of mean GC content of reads. An acceptable mean should be ~40-50% and samples from the same screen should have similar GC content; the SAM samples satisfy those criteria.

Following satisfactory QC results at the sequence level, read count level QC was performed. Read count level QC is based on the sgRNA read counts generated from mapping raw sequence reads to sgRNA sequences thereby assigning each read to one of the 70,290 sgRNAs present in the SAM library (Figure 7.8). The mapping process was conducted by MaGECK with zero tolerance for nucleotide mismatches[380]. Read count level QC included the determination of the percentage of reads successfully mapped to the expected sgRNA sequences, calculation of the number of sgRNAs with 0 read count, and determination of the Gini Index which measures the evenness of the read counts per sgRNA. The SAM screen showed slightly less than ideal 'mapability' with ~37% of reads not mapping to any of the sgRNAs; greater than 65% mapped is deemed optimal. Allowing for tolerance of one nucleotide mismatch during mapping would likely have improved this figure but over 10 million mapped reads for each sample was considered sufficient. The number of missed sgRNAs per sample is well below the maximum threshold of 1%[381]. The Gini index across all samples is elevated compared to the Day 0 sample, which represents the starting point of the screen population. This was expected

and indicated that a selective pressure was exerted on Raji cells by the drugs and the sgRNA-induced phenotypes.

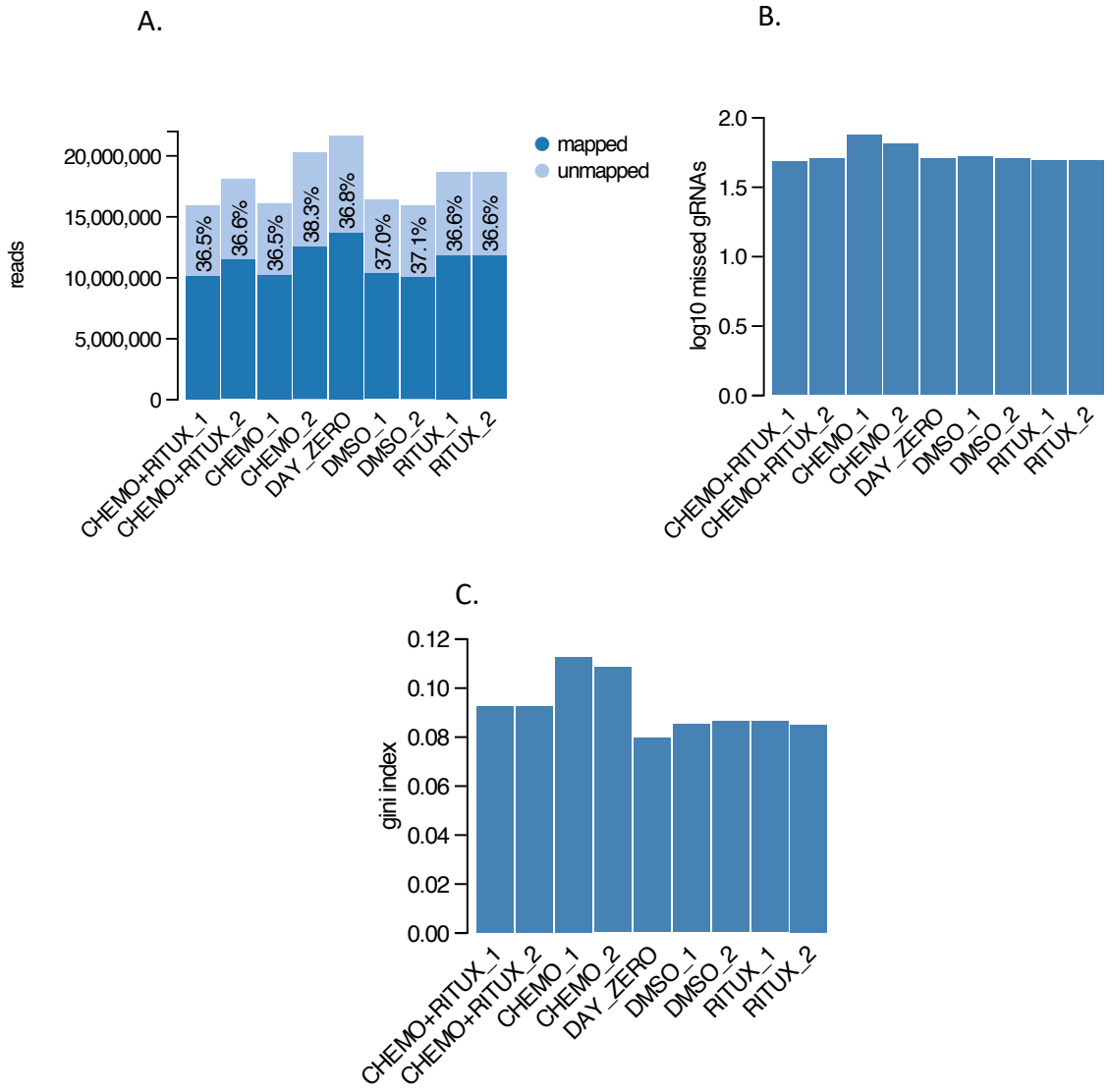


Figure 7.8. Read count level QC of samples from the SAM screen.

(A) Mapability of sequence reads to expected sgRNA guide sequences present in the SAM library. Indicated %s represent unmapped reads (B) Log-transformed counts of sgRNAs with no reads; represents missing sgRNAs from the screen library in each sample. (C) Gini index, which measures inequality between read counts with 1.0 indicating maximum inequality (i.e. all reads mapped to a single sgRNA within the library), 0.0 corresponding to maximum equality with an identical read count for all sgRNAs within the library.

Sample level QC was conducted to analyse the level of consistency between samples (Figure 7.9). Sample level QC determined the distributions of normalized sgRNA read counts, represented by a box plot and a cumulative distribution function (CDF) plot.

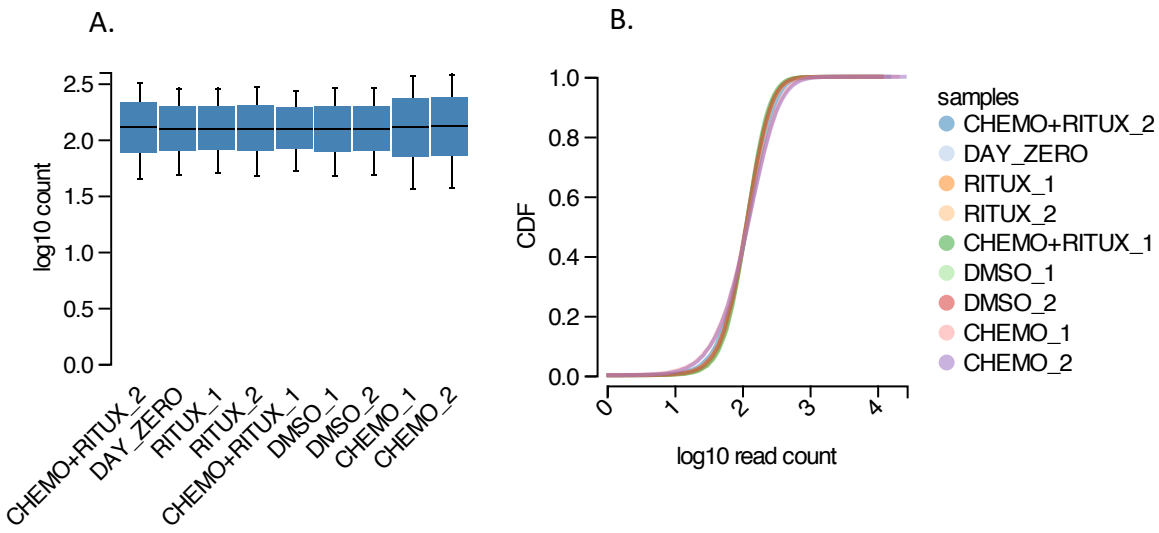


Figure 7.9. Sample level QC of samples from the SAM screen.

(A) Log-transformed read counts of the sgRNAs in a box plot. (C) Log-transformed read counts of the sgRNAs in a CDF plot show relatively equal distribution of log-transformed read counts across samples.

7.2.4 Data analysis of the SAM screen to identify putative Chemotherapy and Rituximab resistance mediators

7.2.4.1 MaGECK-maximum likelihood estimation approach

Several different distinct approaches of data analysis exist to interpret screen data. Traditional methods include RNAi gene enrichment ranking (RIGER) or Redundant siRNA Activity (RSA) analysis[380]. More recent introductions including MaGECK version 1 RRA, which relies on a modified robust ranking aggregation have improved upon these methods. Recently the second edition of MaGECK, MaGECK-VISPR, which uses a maximum likelihood estimation (MLE) approach instead of RRA, was released[381]. Briefly, the key advantage of MaGECK-MLE over MaGECK-RRA (or RIGER, RSA) is its improved capability to accurately compare gene enrichment or dropout across multiple treatment conditions within a single comparison. More specifically, MaGECK-MLE does not rely on a rank-based analysis, allowing for greater accuracy and detailed comparisons of gene selections. Therefore, MaGECK-MLE allows a screener to distinguish 'stronger' enrichment from 'strong' enrichment derived from their actual fold differences instead of their rank differences. This allows MaGECK-MLE to establish more refined clusters of genes that could not be revealed by previous methods. Therefore, MaGECK-MLE was applied to conduct k-means clustering analysis for the SAM screen (Figure 7.10). It appears that the chemotherapy and chemotherapy plus Rituximab samples are more similar to each other whereas the Rituximab only arm has a unique set of enriched and depleted genes.

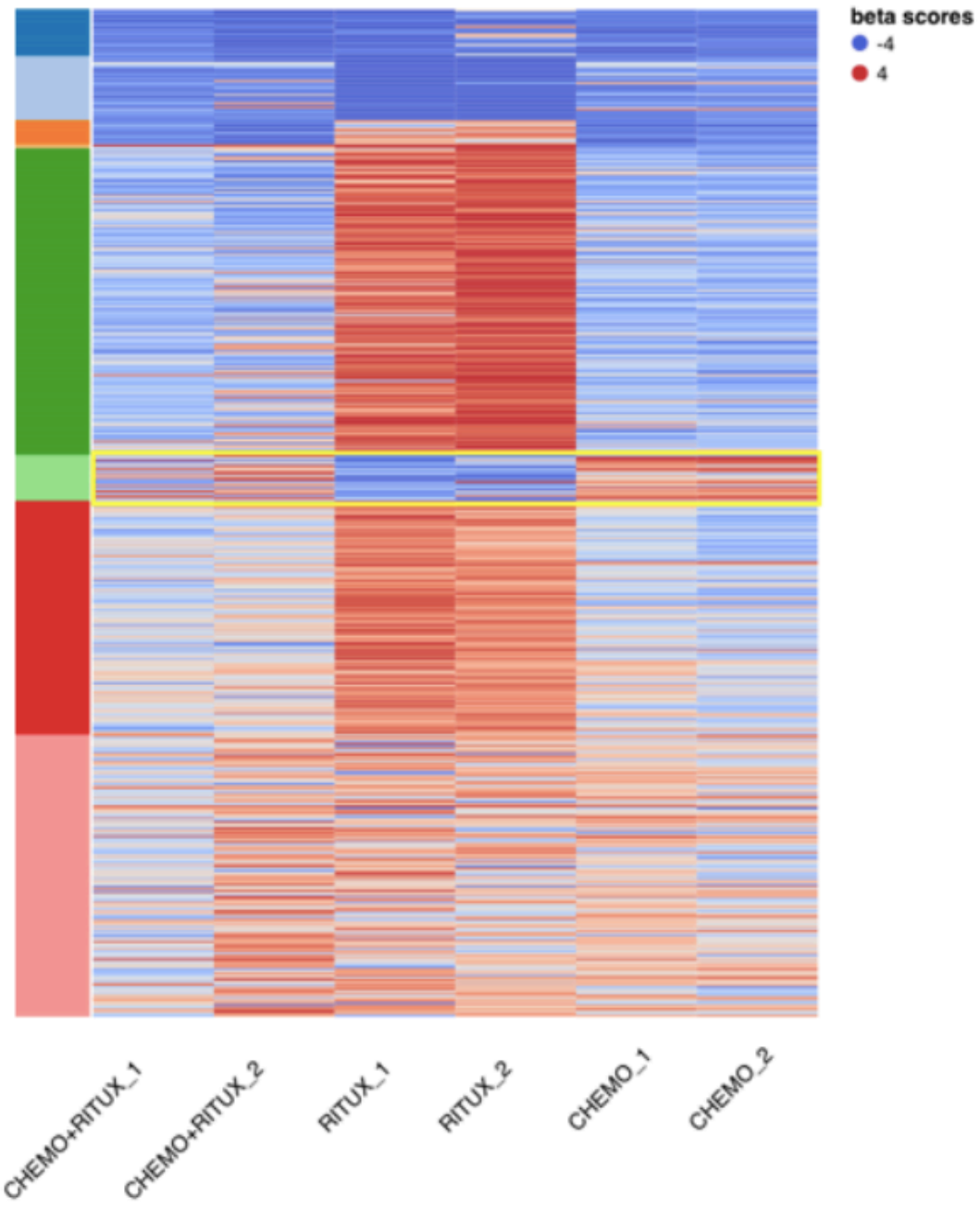


Figure 7.10. k-means clustering views of β scores derived from MaGECK-MLE of conditions in the SAM screens from the top selected genes (k = 8).

The yellow box highlights the cluster of genes selectively enriched in chemotherapy and chemotherapy plus Rituximab treated samples. Red indicates enrichment whilst blue shows depletion.

In addition, employing a default FDR of < 0.5, MaGECK-MLE analysis of the OV screen identified 16 candidate genes for the chemotherapy, 12 for the chemotherapy plus Rituximab and 17 for the Rituximab only arms (Table 7.2)

Chemotherapy	Chemotherapy + Rituximab	Rituximab
ABCC1	ABCC1	IRAK1
ABCB1	ABCB1	SLITRK4
PIM1	PIM1	PPP5D1
ADAM9	ADAM9	GPR65
JARID2	JARID2	FBXO36
PEG10	PEG10	ITGB1
DDB2	DDB2	ATG10
DKK3	DKK3	LIN7B
MAD2L2	MAD2L2	CITED2
NFATC4	NFATC4	RAC2
PDCD1	ALCAM	ALCAM
BRCC3	ASAP3	ASAP3
MAP3K4		MAPK13
IKZF2		ZC4H2
ARGEF1		IL36B
CBX3		RAB12
		ROM1

Table 7.2. MLE-derived 'hits' with a FDR <0.5

Red identifies hits common to both the chemotherapy and the chemotherapy plus Rituximab arms. Blue identifies hits common to both the Rituximab and chemotherapy plus Rituximab arms. Black identifies hits present in one arm only.

Perhaps as expected, ABCB1 and ABCC1 are by far the most significant hits detected in both the chemotherapy and chemotherapy plus Rituximab arms with both a *p*-value and FDR of 0 and several hundred-fold enrichment, implying ABC transporters are the most efficient resistance mechanism against multi-agent chemotherapy. In the chemotherapy + Rituximab arm there is significant overlap with hits in the chemotherapy only arm. These include the ABC transporters, anti-apoptotic gene PIM1, NFATC4 - a member of the NFATC transcription complex - and genes involved in DNA damage response including BRCC3 and DDB2. Despite the clear overlap, the chemotherapy plus Rituximab arm has less hits than the Chemotherapy only arm. This may be due to the introduction of a fourth drug with a completely different mechanism of action making it more difficult for MAGECK to call significance. Rituximab alone has a distinct set of hits involved in TLR signalling, the MAP Kinase pathway and signalling by Rho GTPases. It shares some common hits with the chemotherapy plus Rituximab arm including ALCAM and ASAP3 but overall, as suggested by the k means clustering shown in Figure 7.10, it appears that the dominant resistant mechanisms of the chemotherapy plus Rituximab arm are driven by the traditional chemotherapeutic drugs.

7.2.4.2 Fold difference and weighted average analysis

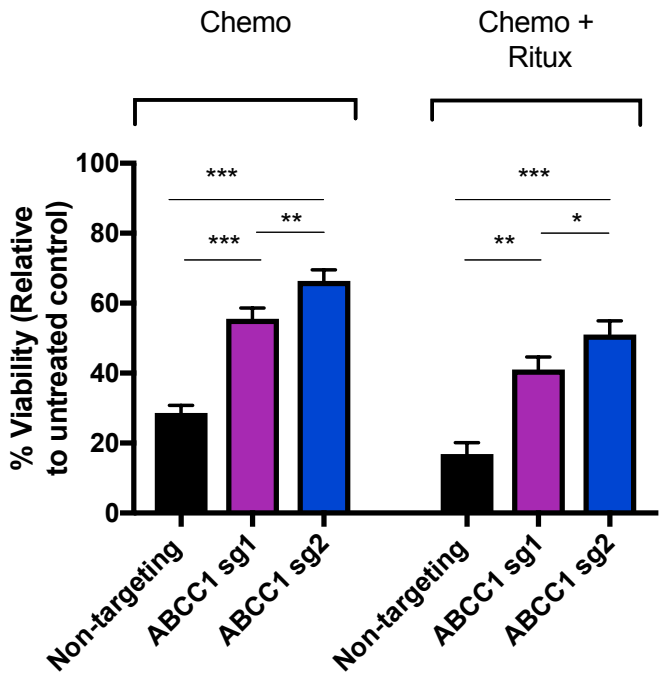
Analysis of the enriched genes reported by MaGECK-VISPR highlighted that that there would be a significantly higher number of identified candidate genes if *p*-values associated with enriched genes were used to compile the lists. Essentially, the MLE approach may have been too stringent. Hence, an alternative, simplified analysis method based on the calculation and comparison of fold differences (FD) of sgRNAs was applied following consultation with bioinformaticians. This approach required manual analysis of the sgRNA counts and did not necessarily require the computing power of bioinformatics scripts. Briefly, the weighted average of the sgRNA read counts between biological replicates were calculated first for all of the sgRNAs present in the library. The derived weighted average read counts of the sgRNAs from drug-treatment samples were compared against those from the DMSO-exposed samples to identify sgRNAs that were preferentially enriched under the various drug treatment conditions. Within each treatment condition, an arbitrary threshold (>1.5 FD) was set and the genes with multiple sgRNAs exceeding the FD threshold were considered to be significantly enriched. For the purpose of this study, due to the use of multiple drugs within treatment arms, the threshold was increased to 5 fold for extra stringency instead of the 1.5 FD suggested in the above analysis method. This method resulted in 72 hits in the Chemotherapy only arm, 55 hits in the Chemotherapy plus Rituximab arm and 92 hits in the Rituximab arm (Appendix 1).

7.2.5 Individual validation of candidates identified from the SAM screen

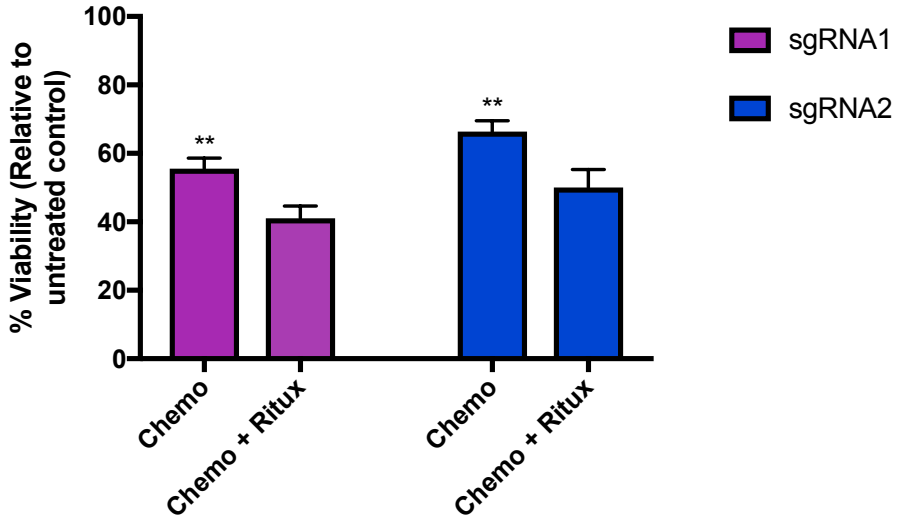
In order to validate hits from the SAM screen, 21 genes were selected for analysis and functionally validated by individually transducing Raji cells with two sgRNAs per gene. In some cases, where only one sgRNA out of three had given significant enrichment, an alternative sgRNA, not in the screen library was chosen. This approach was taken based on the fact that since the design of the CRISPR activation library v1, more specific and efficient sgRNAs for many of the genes have been identified. Indeed, the Feng Zhang lab provides a tool, in which guide RNAs for each gene are sorted in order of specificity, based on methods described in their publication[375]. Examination of the sgRNAs employed in this study with the tool highlighted that for most of the genes identified by the screen, that only showed enrichment for one of the three sgRNA, there were several more specific predicted sgRNAs. In fact, the three sgRNAs present in the screen library were often the lowest ranked sgRNAs. The sgRNA that had resulted in a positive hit in this screen was therefore selected for validation but with a second sgRNA replaced and selected using the tool, as the one ranked as most specific for that gene. The selected two sgRNAs per candidate gene were then cloned and packaged into lentivirus particles to produce 43 stable cell lines (42 sgRNAs targeting candidate genes and 1 non-targeting control). In addition, validation was carried out on some hits that were not identified by the original analysis method (i.e. with a FDR<0.5) but were instead identified by the FD method, in order to determine if this analysis method is reliable. Genes identified by the FD method that were included in validation are marked with # in the figures. The established cell lines were treated with drugs at the ED value used in the screen for three days, after which cell viabilities were measured and compared to NT sgRNA transduced control cells. Increased ABCC1 expression induced by two alternative sgRNAs resulted in a significant increase in cell viability following treatment with both multiagent chemotherapy and chemotherapy plus Rituximab compared to control NT sgRNA (Figure 7.11A). The fact that the viability of cells is significantly lower compared to the untreated control in the chemotherapy plus Rituximab treated ABCC1 expressing cells than the chemotherapy only treated cells (Figure 7.11B), suggests that the increased expression of ABCC1 may not be driving resistance to Rituximab. Analysis of ectopic levels of ABCC1 expression due to SAM activity highlighted that sgRNA 2 results in a greater fold change

induction of ABCC1 mRNA than sgRNA 1, likely explaining the significantly increased viability of sgRNA2 derived cell lines (Figure 7.11C).

A.



B.



C.

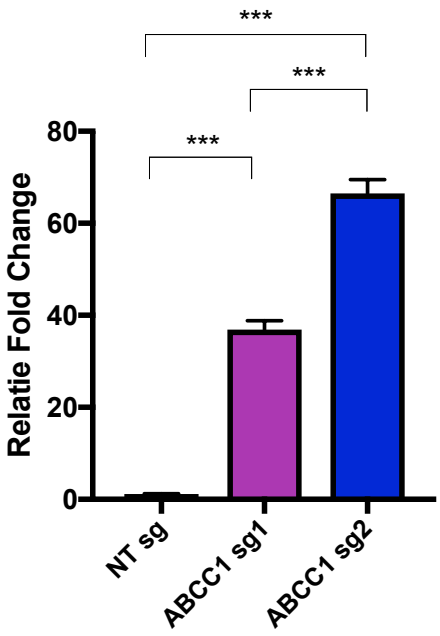


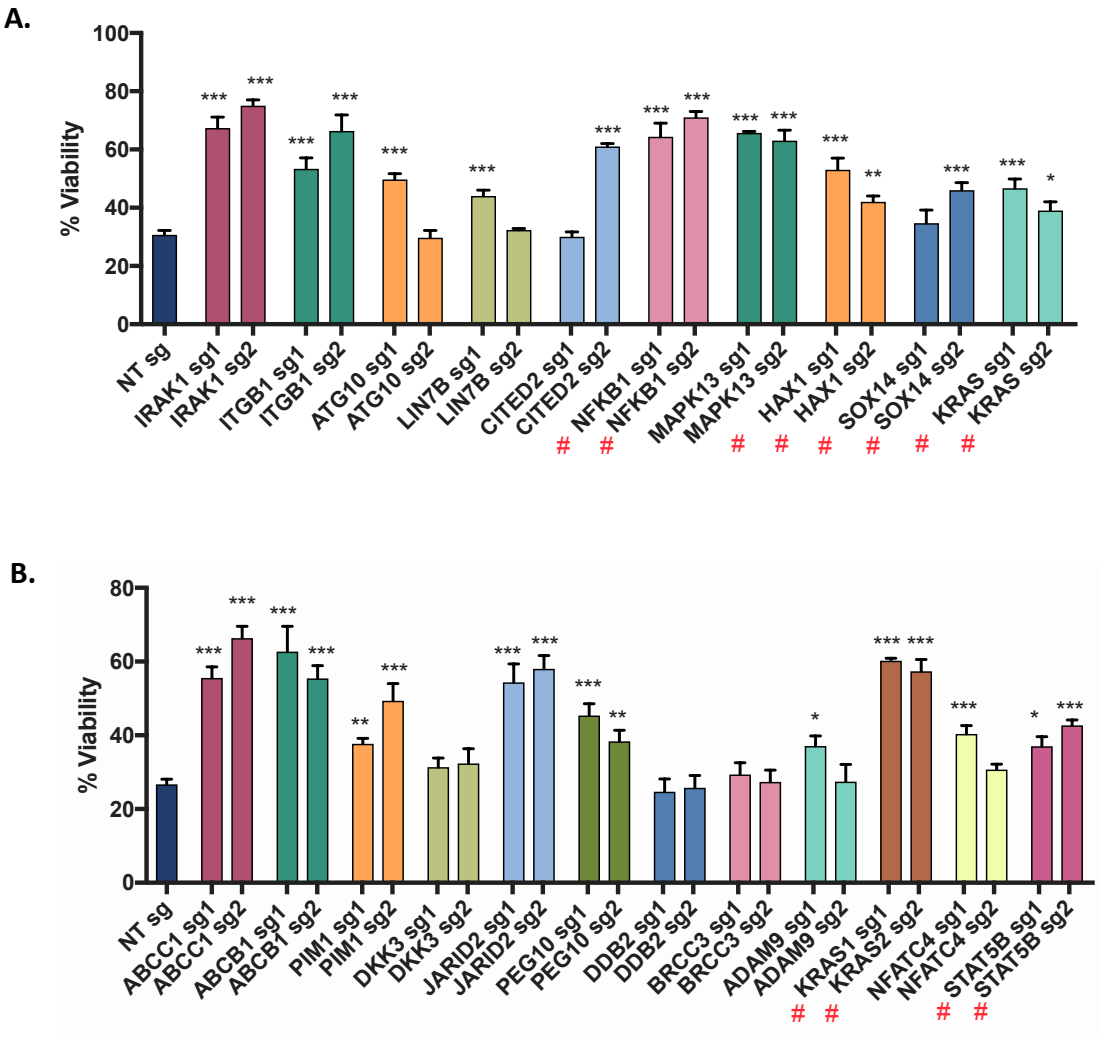
Figure 7.11. Ectopic expression of ABCC1 by CRISPR SAM activity induces resistance to multiagent chemotherapy

(A). sgRNA transduced cell lines were treated with either chemotherapy or chemotherapy plus Rituximab for 72 hours at the concentration utilised in the screen (ED70 and ED85 respectively). Cell viability was determined by cell titre blue assay. Data represent the means and standard deviations of technical triplicates.

(B). Viability of both ABCC1 sgRNA1 and sgRNA2 transduced cell lines compared to untreated control cells is less when treated with chemotherapy plus Rituximab compared to chemotherapy alone.

*(C). RT-qPCR to determine the level of overexpression induced by SAM screen sgRNAs. Fold-differences were determined relative to NT sgRNA for both of the sgRNAs designed to target the TSS of the ABCC1 gene. RT-qPCRs were performed from RNA extracted 5 days after sgRNA transduction and conducted in technical triplicates, means and standard deviations are shown. Chemo = chemotherapy, Chemo + Ritux = chemotherapy plus Rituximab, sg1 = ABCC1 sgRNA1, sg2 = ABCC1 sgRNA2. **p<0.01, ***p<0.001*

This validation process was repeated for each of the selected sgRNAs and the viability of the transduced cell lines following treatment calculated (Figure 7.12). After completing the validation process across all three arms of the screen, a trend was seen whereby the number of successfully validated sgRNAs and the significance level was greater in the single agent Rituximab arm than in the multi-agent arms. Of the sgRNAs tested, 80% conferred resistance to Rituximab alone, including 7 out of 8 sgRNAs that were identified by the FD method. For chemotherapy, the percentage of validated sgRNAs was 67% while 63% of sgRNAs were validated in the chemotherapy plus Rituximab arm. This implies that combination treatment may result in the reporting of significant hits that do not confer resistance to all drugs. These are not technically false negatives as resistance and enrichment did occur but only as a result of an sgRNA conferring resistance to some of the agents used.



C.

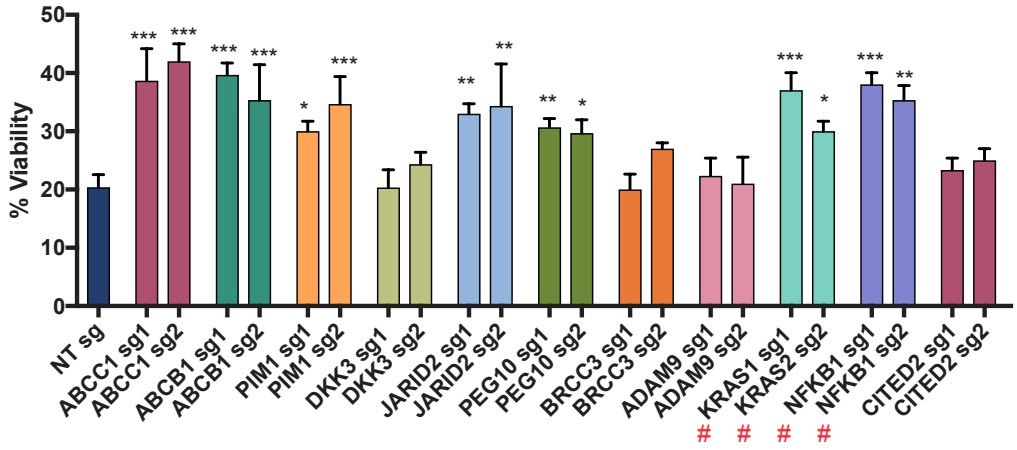
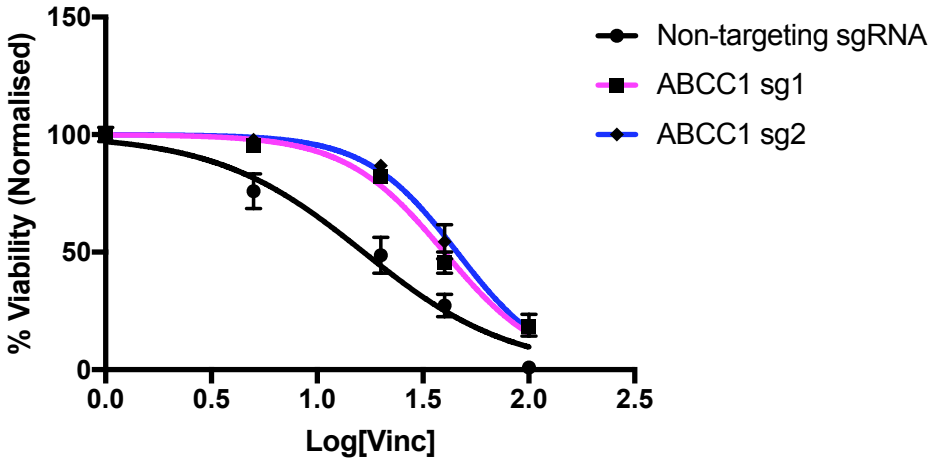


Figure 7.12. Viability of sgRNA transduced cells following drug treatment

*sgRNA transduced cell lines were treated with either Rituximab (A) chemotherapy (B) or chemotherapy plus Rituximab (C) for 72 hours at the concentration utilised in the screen (ED70, ED70 and ED85 respectively). Cell viability was determined by the cell titre blue assay. Data represent the means and standard deviations of technical triplicates *p<0.05, **p<0.01, ***p<0.001. # marks sgRNAs identified by the FD method.*

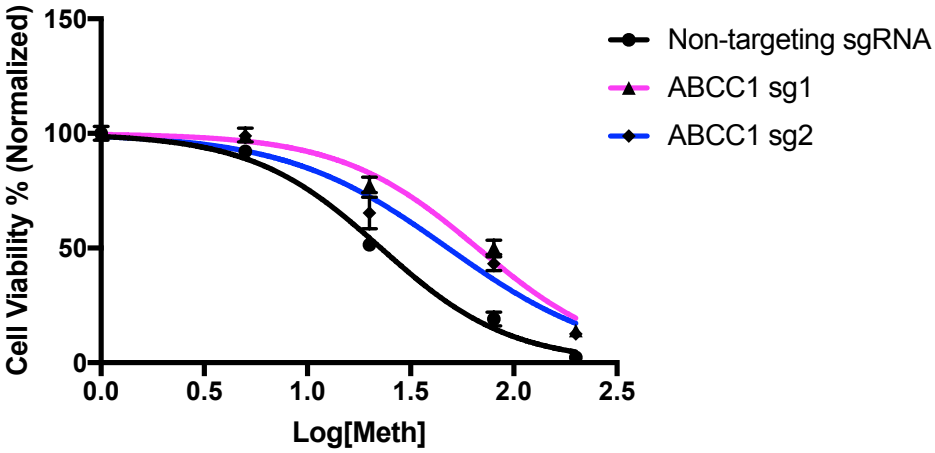
In order to determine if the increased viability induced by sgRNA transduction in many cases was due to increased resistance to all drugs or just some, the established cell lines were treated with a range of concentrations of single agent drug for 3 days, after which the cell viabilities were measured and the ED50 was established for each pair of sgRNA transduced cell lines as well as the NT sgRNA cell line. An example of the data for ABCC1 for doxorubicin, vincristine, methotrexate and Rituximab is shown in Figure 7.13 along with the determined ED50 values.

Vincristine

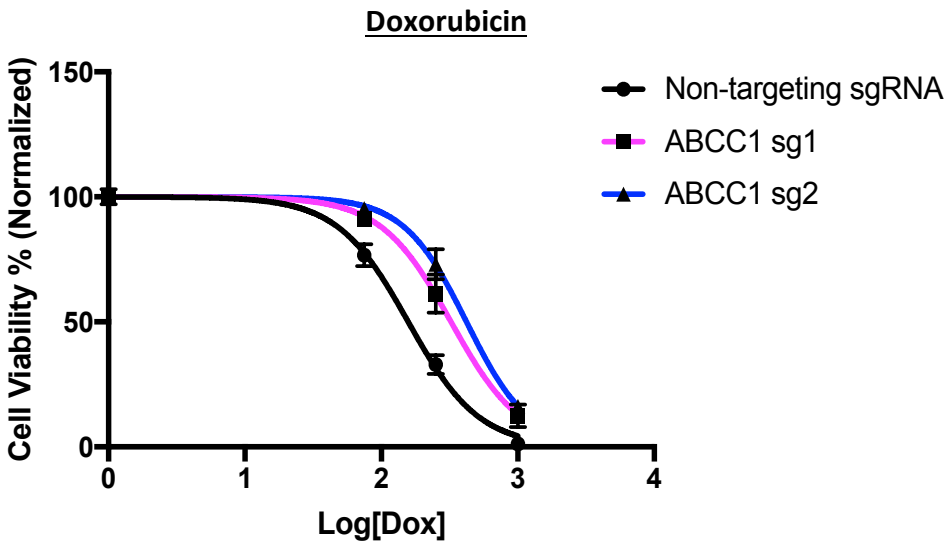


Non-targeting sg	ABCC1 sg1	ABCC1 sg2
1.6nM	3.9nM	4.5nM

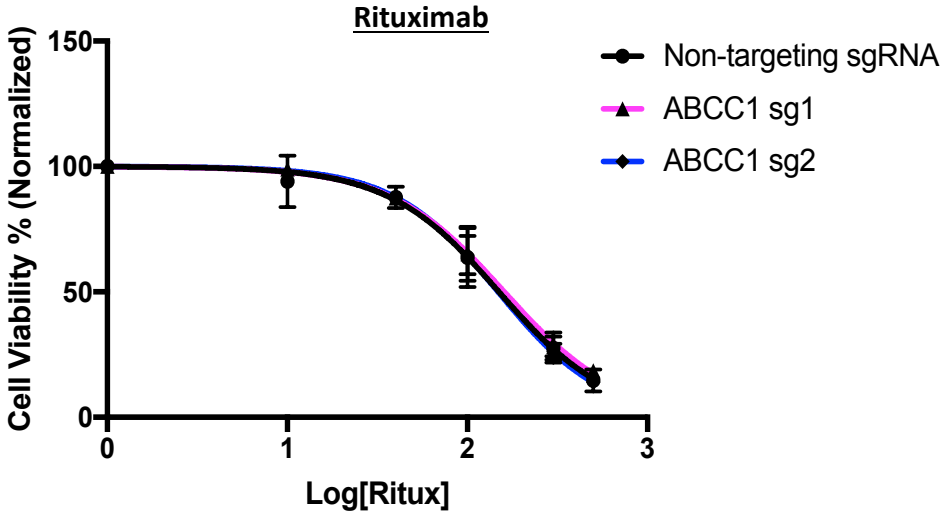
Methotrexate



Non-targeting sg	ABCC1 sg1	ABCC1 sg2
22nM	66nM	49nM



Non-targeting sg	ABCC1 sg1	ABCC1 sg2
157nM	324nM	420nM



Non-targeting sg	ABCC1 sg1	ABCC1 sg2
965nM	1002nM	940nnM

Figure 7.13. Overexpression of ABCC1 induces resistance to chemotherapy but not Rituximab.

Cell viability was normalized to Raji cells treated with vehicle control alone. Non-linear regression plots display mean viabilities and the 95% confidence interval of technical triplicates with ED50s listed below. All shifts in ED50 values for doxorubicin, methotrexate and vincristine were determined to be significant ($p < 0.001$) by one-way ANOVA analysis. Rituximab was not significant (Table 7.3).

These data highlight that although ABCC1 was a significant hit in the chemotherapy plus Rituximab arm of the screen, the increase in cell viability conferred by over expression of this gene was only due to increased resistance to the three chemotherapy drugs. There was no significant ED50 shift for the ABCC1-transduced cell lines compared to the NT sgRNA transduced cell line highlighting that ABCC1 does not confer resistance to Rituximab. Similar results were seen throughout the single drug validation process. Table 7.3 shows the statistical significance levels of the ED50 shift conferred by each sgRNA-transduced cell line treated with single agents and highlights that PIM1, PEG10 and KRAS caused a significant ED50 shift for all 4 drugs, NFKB1 confers resistance to two chemotherapy drugs and Rituximab, while ABCB1, ABCC1 and JARID2 confer resistance to each of the chemotherapies but not Rituximab.

sgRNA ID	Doxorubicin	Methotrexate	Vincristine	Rituximab
ABCC1 sg1	***	***	***	ns
ABCC1 sg2	***	***	***	ns
ABCB1 sg1	***	**	***	ns
ABCB1 sg2	***	***	**	ns
PIM1 sg1	**	**	**	*
PIM1 sg2	***	***	*	**
JARID2 sg1	***	***	**	ns
JARID2 sg2	**	***	***	ns
NFKB1 sg1	**	ns	***	*
NFKB1 sg2	**	ns	**	**
KRAS sg1	*	*	**	**
KRAS sg2	***	**	**	ns
PEG10 sg1	*	**	**	**
PEG10 sg2	**	**	*	*
MAPK13	n/a	n/a	n/a	**
MAPK13	n/a	n/a	n/a	*
IRAK1 sg1	n/a	n/a	n/a	***
IRAK1 sg2	n/a	n/a	n/a	***
ITGB1 sg1	n/a	n/a	n/a	**
ITGB1 sg2	n/a	n/a	n/a	**

Table 7.3. Significance levels of ED50 shifts for the indicated sgRNA-transduced cells exposed to single drugs

*One-way ANOVA tests were performed for all conditions to determine the significance of the shifts in drug sensitivities compared to that observed in NT sgRNA transduced cells. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns – not significant. n/a – drug not tested.*

7.2.6 Different pathways drive resistance to chemotherapy and Rituximab *in vitro*

The previous section highlighted the successful validation of sgRNAs identified by the FD method indicating that, as expected the FDR threshold is overly stringent and the second method of validation is reliable and effective at identifying numerous genes conferring resistance to the drugs used in this screen. The 92 genes identified as having greater than 5 fold enrichment in the Rituximab arm were subjected to GSEA and overlap with gene sets in the Broad Institute MSigDB database was determined. Hallmark and KEGG gene sets were included in the output which identified some interesting pathways

implicated in conferring resistance to Rituximab including MAP Kinase and Toll-like receptor signalling as well as genes associated with up regulation of chemokine signalling (Figure 7.14). Similarly, the 72 identified hits in the chemotherapy arm were also subjected to GSEA to identify dominant molecular signatures in the hit list. Importantly, it should be remembered that at the validation stage a lower number of hits in the chemotherapy arm were successfully validated upon treatment with all drugs (67%), so it is likely some hits in this list are conferring resistance to only 1 or 2 of the three chemotherapeutics. Nonetheless, this gives us an overview of the enriched signatures in response to these chemotherapeutic drugs. Interestingly, a different set of pathways to Rituximab were identified as significant when these hits were compared with gene sets and PI3K signalling was one the most significant pathways in the chemotherapy treatment arm. This suggests that Rituximab, although essentially acting as a chemotherapeutic in the absence of immune components, has a different set of resistance mechanisms to traditional chemotherapeutics.

A.

Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	k/K	p-value	FDR q-value
KEGG_MAPK_SIGNALING_PATHWAY [267]	MAPK signaling pathway	12		$4.68\ e^{-13}$	$1.1\ e^{-10}$
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHATHWAY [102]	Toll-like receptor signaling pathway	8		$5.02\ e^{-11}$	$5.93\ e^{-9}$
KEGG_CHEMOKINE_SIGNALING_PATHWAY [190]	Chemokine signaling pathway	9		$2.77\ e^{-10}$	$2.18\ e^{-8}$
KEGG_FC_GAMMA_R_MEDIATED_PHAGOCYTOSIS [97]	Fc gamma R-mediated phagocytosis	7		$1.53\ e^{-9}$	$9.03\ e^{-8}$
KEGG_PANCREATIC_CANCER [70]	Pancreatic cancer	6		$8.35\ e^{-9}$	$3.61\ e^{-7}$
KEGG_NEUROTROPHIN_SIGNALING_PATHWAY [126]	Neurotrophin signaling pathway	7		$9.58\ e^{-9}$	$3.61\ e^{-7}$
KEGG_AXON_GUIDANCE [129]	Axon guidance	7		$1.13\ e^{-8}$	$3.61\ e^{-7}$
KEGG_B_CELL_RECEPTOR_SIGNALING_PATHWAY [75]	B cell receptor signaling pathway	6		$1.27\ e^{-8}$	$3.61\ e^{-7}$
KEGG_VEGF_SIGNALING_PATHWAY [76]	VEGF signaling pathway	6		$1.38\ e^{-8}$	$3.61\ e^{-7}$
KEGG_FC_EPSILON_RI_SIGNALING_PATHWAY [79]	Fc epsilon RI signaling pathway	6		$1.74\ e^{-8}$	$4.11\ e^{-7}$

B.

Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	k/K	p-value	FDR q-value
KEGG_PATHWAYS_IN_CANCER [328]	Pathways in cancer	14		$7.32\ e^{-17}$	$1.73\ e^{-14}$
HALLMARK_PI3K_AKT_MTOR_SIGNALING [105]	Genes up-regulated by activation of the PI3K/AKT/mTOR pathway.	9		$6.04\ e^{-14}$	$7.13\ e^{-12}$
KEGG_ACUTE_MYELOID_LEUKEMIA [60]	Acute myeloid leukemia	7		$4.56\ e^{-12}$	$3.59\ e^{-10}$
KEGG_PANCREATIC_CANCER [70]	Pancreatic cancer	7		$1.4\ e^{-11}$	$7.45\ e^{-10}$
KEGG_NEUROTROPHIN_SIGNALING_PATHWAY [126]	Neurotrophin signaling pathway	8		$1.85\ e^{-11}$	$7.45\ e^{-10}$
KEGG_CHRONIC_MYELOID_LEUKEMIA [73]	Chronic myeloid leukemia	7		$1.89\ e^{-11}$	$7.45\ e^{-10}$
KEGG_B_CELL_RECEPTOR_SIGNALING_PATHWAY [75]	B cell receptor signaling pathway	7		$2.3\ e^{-11}$	$7.76\ e^{-10}$
KEGG_PROSTATE_CANCER [89]	Prostate cancer	7		$7.86\ e^{-11}$	$2.32\ e^{-9}$
KEGG_NON_SMALL_CELL_LUNG_CANCER [54]	Non-small cell lung cancer	6		$2.24\ e^{-10}$	$5.87\ e^{-9}$
KEGG_T_CELL_RECEPTOR_SIGNALING_PATHWAY [108]	T cell receptor signaling pathway	7		$3.11\ e^{-10}$	$7.34\ e^{-9}$

Figure 7.14. GSEA analysis of hits identified as conferring resistance to Rituximab and chemotherapy in the SAM screen.

Gene sets in MSigDB, specifically hallmark gene sets and KEGG gene sets, were compared to the hit lists. The top enriched pathways (FDR < $0.1X10^{-5}$) are shown for the Rituximab candidate list (A) and the chemotherapy candidate list (B).

7.2.7 Clinical relevance of the identified candidates

An important consideration when utilising prospective screen based studies in research is how translationally relevant the findings will be. For example if an identified candidate confers resistance upon ectopic expression *in vitro* but is never expressed in patient tumours the translational potential is limited. In order to determine if some of the most significant hits in this SAM screen may be clinically relevant targets, endogenous levels of expression of the genes were examined in the RNA-seq PDX data generated in chapter 6. Kaplan-Meier EFS analysis based on expression level of the candidate genes identified, was also produced from larger published patient cohorts with matched clinical data. All genes that were included in the validation process were expressed in our untreated

PDXs with log-transformed expression levels of example genes shown in Figure 7.15. In addition, using PRECOG (PREdiction of Clinical Outcomes from Genomic Profiles)[571], associations between gene expression and outcome were investigated by Kaplan-Meier analysis of treatment naïve BL datasets from the Hummel and Dave studies[58,572] (Figure 7.16). Exact patient age was not available for the Hummel cohort but 50% of patients in the Dave study were under 18 years of age. IRAK1 expression has the highest association with reduced OS (Figure 7.16). Other identified hits from our study that did not reach statistical significance, likely due to the limited number of patients, but showed a trend towards worse OS with increased expression, include PEG10, PIM1 and SOX14.

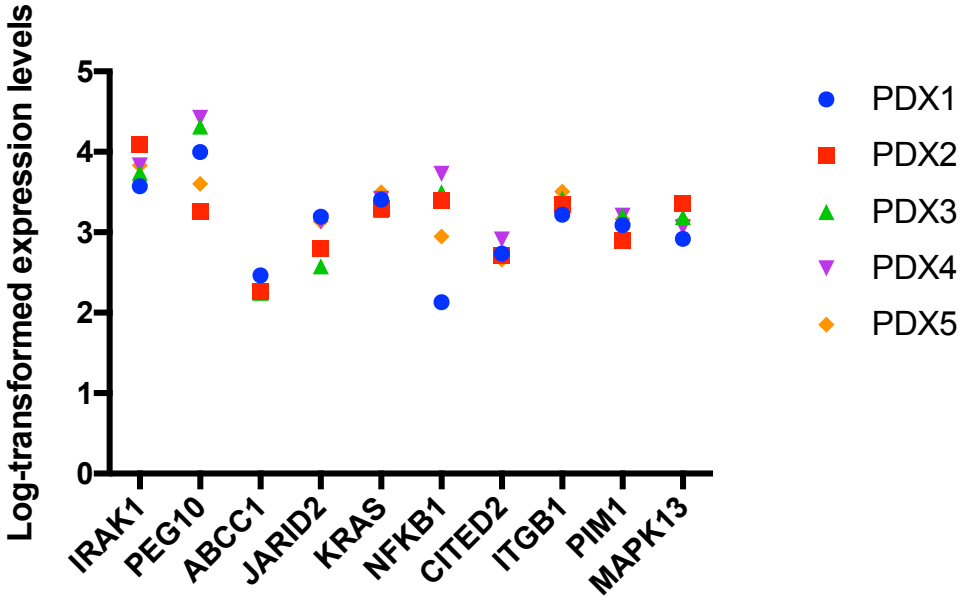


Figure 7.15. Candidate genes conferring resistance to chemotherapy and Rituximab are expressed endogenously in patient-derived PDX samples.

Log-transformed read counts for each gene in all five PDX samples are displayed. Read counts were determined following mapping of RNA-seq data to the genome.

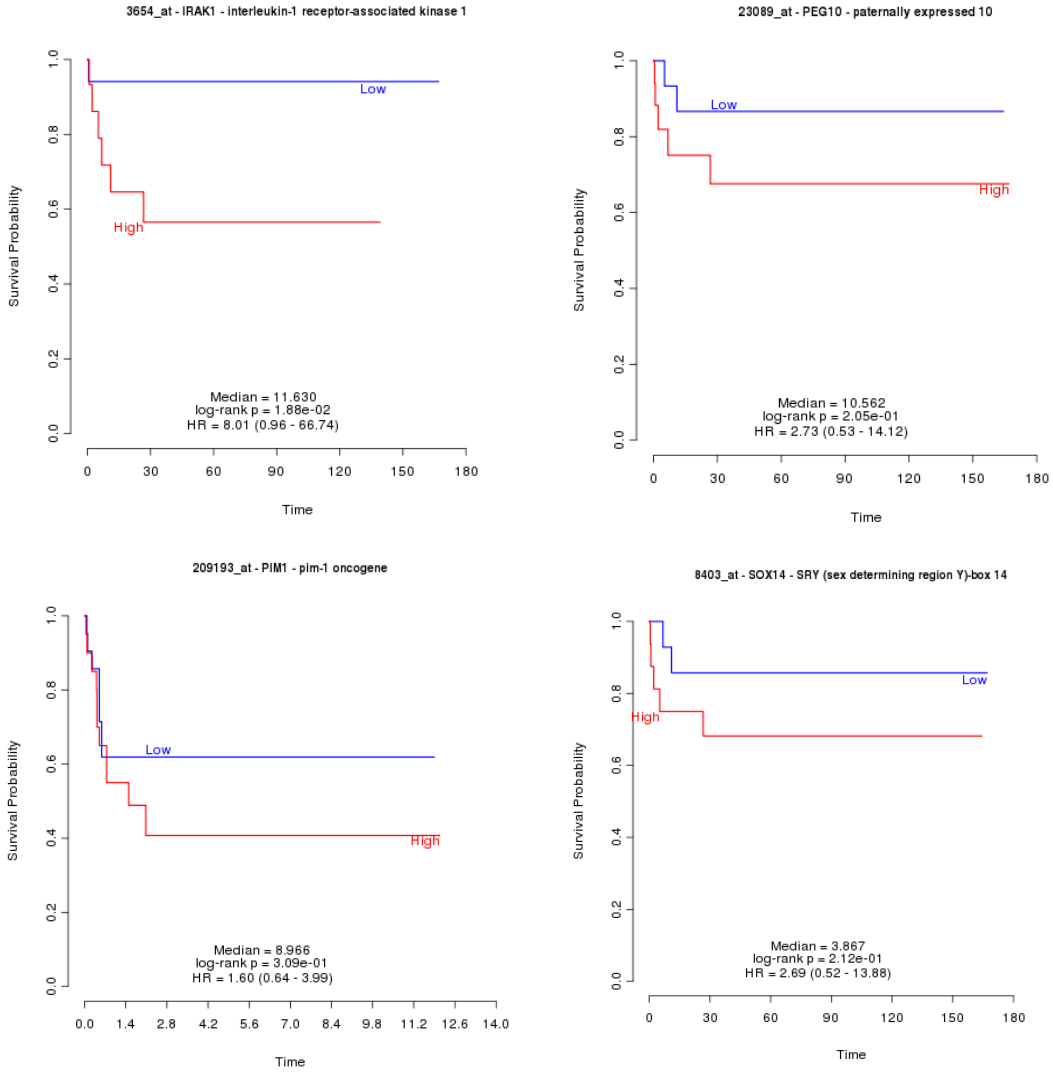


Figure 7.16. Kaplan-Meier OS trends align with candidates identified from the SAM screen.

PRECOG Kaplan-Meier OS analysis was carried out in BL datasets by gene expression. Only IRAK1 shows a statistically significant worse OS (log-rank test, p=0.018) with increased expression but other genes show a similar trend.

7.3 Discussion

The SAM screen with chemotherapy, Rituximab or chemotherapy plus Rituximab treatment arms produced a number of interesting observations. Initial validation of the CRISPR/dCas9-VP64 induced overexpression phenotype highlighted that not every gene in our sample gene set was significantly upregulated upon targeted sgRNA transduction in the Raji cell line. Furthermore, in some cases the induced expression level varied significantly between the two tested sgRNAs. This result was also noted during validation of SAM activity by other researchers in the lab, using different cancer type cell lines. In one case a trend was determined whereby lower levels of endogenous (unmodified) expression level of the target genes were associated with more potent induction of SAM-mediated overexpression of the same genes. However, this trend was not seen in the Raji cell line and any failure of induction of overexpression is likely due to the level of specificity and efficiency of individual sgRNAs within the library. Nonetheless, 11 out of 15 of the genes tested showed a greater than 10 fold increase in mRNA levels in response to the SAM system and hence a large majority of the 23,430 genes targeted in this screen should have successful induction of overexpression.

Before conducting any gene-based analysis to identify potential resistance mechanisms, three layers of QC analysis were performed on the data with MaGECK-VISPR. Specifically, QC of all raw fastq files derived from the screen samples were conducted at the sequence, read count and sample levels. Following verification that all levels of QC exceeded the minimum threshold requirements, a concise list of genes was derived for the three treatment arms based on FDR values using the MaGECK-MLE approach. The most significant hits in the chemotherapy containing arms were ABC transporters which is somewhat expected but also an interesting observation and a positive control for the successful conduction of the screen. The identification of strongly enriched hits such as IRAK1 in the Rituximab arm is also particularly interesting as the exact mechanisms of action of this drug *in vitro,* in the absence of ADCC and CMC, are poorly understood.

During the validation stage the complexity introduced by multiagent treatment was highlighted and likely explains the relatively short list of hits produced. While certain sgRNAs increased cell viability following chemotherapy plus Rituximab treatment,

investigation of the ED50 shift produced by these sgRNAs for each single agent highlighted that some sgRNAs only conferred resistance to certain agents. For example, ABCC1 which is a drug transporter known to have doxorubicin, methotrexate and vincristine as substrates was the most significant hit for both the chemotherapy and chemotherapy plus Rituximab arms with a *p* value and FDR of zero. However, single agent dose response curves of the ABCC1 sgRNA transduced cell lines compared to the non-targeting sgRNA transduced cell line verified that ABCC1 does not confer resistance to Rituximab and the increased viability was due only to the significant resistance ABCC1 induces to the three other drugs. It is therefore clear that when conducting multidrug treatment screens, the validation process is extremely important and also significantly more laborious than for single agent screens. Furthermore, ensuring the correct balance of each drug is also a challenge. Significant effort was made to determine optimum doses of each drug to ensure contribution of each agent to the induction of resistance. Using the bliss independence model concentrations of vincristine, doxorubicin and methotrexate were identified that resulted in minor synergy. For the addition of Rituximab it was decided to keep the chemotherapy concentrations the same to have a direct comparison between the two arms, with the addition of Rituximab the only variable. This resulted in minor synergy and an ED value of 85. An unexpected observation over the course of the screen was the poor survival of cells in the chemotherapy plus rituximab arm, beyond the expected effect of the ED85 dose. This highlighted that the dosing may have been incorrect or that with 4 drugs in total, at relatively low individual concentrations and with diverse mechanisms of action, there was not enough selective pressure from any to drive a resistant phenotype.

Despite the complexity of a multiagent screen and the requirement for single agent validation, hits that were identified in both the chemotherapy and chemotherapy plus Rituximab arms, including PIM1 and PEG10, were successfully validated and shown to confer resistance to all 4 drugs. PIM1, a serine/threonine protein kinase has been associated with BL in previous studies and is enhances the rate of occurrence of c-Myc-induced lymphomas, by blocking genotoxic stress-induced apoptosis[573]. It is mainly involved in the development of hematopoietic malignancies, including BL by stimulating cell survival and proliferation and inhibiting apoptosis through phosphorylation and

subsequent inactivation of pro-apoptotic proteins such as Bad[466,574]. Interestingly, PIM1 expression was shown to be increased in the SP of three of our PDXs, along with other anti-apoptotic proteins, highlighting the significance of this survival pathway in BL and its potential role in driving the stemness phenotype. PEG10 has also been reported in BL before but to a lesser extent; it is a c-MYC target gene in cancer cells and interestingly is located on chromosome 7q, amplifications of which have been associated with an adverse prognosis in BL[67,575,576]. Similarly, PEG10 has been reported to promote progression, proliferation and metastasis of numerous cancers including prostate, breast and lung[577–579]. In CLL, PEG10 is significantly overexpressed in patients and inhibition leads to apoptosis of cells[579]. Furthermore, in BL, PEG10 promotes the migration of cells by up-regulating the expression of matrix metalloproteinases and inhibition of expression again resulted in apoptosis[580]. JARID2 is also a noteworthy gene based on the significant increased resistance it conferred against all three chemotherapy drugs. JARID2 is a cofactor of PRC2 and based on studies involving its relationship with EZH2, its methylation fine-tunes PRC2 activity depending on the chromatin context[581]. Aberrant expression of JARID2 has been reported in several cancers and in ovarian cancer its knockdown inhibits proliferation and invasion through the PI3K/AKT signalling pathway[582]. Interestingly it has been shown to be essential for the maintenance of TICs in bladder cancer[583].

Validation of the hits from the Rituximab arm of the screen was more straight forward and yielded some interesting results. The most significant hit was IRAK1; the IRAKs are key mediators of TLR and interleukin-1 receptor (IL1R) signalling processes. TLR/IL1R-mediated signalling controls diverse cellular processes including inflammation, apoptosis, and cellular differentiation[584]. TLR/IL1R signalling is achieved through differential recruitment of adaptor molecules such as MyD88, Mal/TIRAP, TRIF, and TRAM[585]. These adaptors function in the subsequent recruitment and activation of IRAK family kinases. Cancer-specific dependencies on IRAK signalling became evident following the discovery of oncogenically active MyD88 mutations in ABC DLBCL[586]. Notably, in a large set of tumour biopsies, sequence analysis of the MyD88 coding region revealed that 29% of ABC DLBCL tumours harbour the L265P single amino acid substitution within the MyD88 TIR domain. However, this mutation is absent in GC

DLBCL and BL. Interestingly, however, a recently published comparative genomic expression signature study in pBL identified potential new signal transduction pathways that are involved in pBL. These mainly include TLR, JAK-STAT and MAPK signalling and in particular IRAK1 was shown to be overexpressed 34 fold in this pBL cohort making the discovery that expression of this gene can confer resistance to Rituximab particularly interesting as we move into the era of Rituximab treatment for B-NHL in children[62]. Of course it remains to be investigated if this mechanism of resistance is relevant in the context of ADCC and CMC.

The second method of analysis using the FD method was based on manual analysis of read counts and determination of fold change compared to a DMSO control. It was decided to include this alternative method based on the observation that there were a large number of hits with extremely low *p* values that fell outside the FDR threshold set by MaGECK-VISPR. It was therefore likely that a significant number of positive hits were being omitted from analysis. For the purpose of this study a conservative minimum fold change of 5 was chosen which yielded a significant number of hits in each treatment arm. Most importantly, hits from this method were also included in the validation stage and were successfully validated. These included NFKB1 and KRAS which were both outside the FDR threshold but induced resistance to chemotherapeutic drugs and Rituximab. The validation of genes that did not appear in the original FDR-compiled hit list is an encouraging indicator of the degree of accuracy of the hits derived from the FD method. Furthermore, enrichment screens are generally considered to be less rigorous and more flexible in the data analysis phase compared to negative selection ('dropout') screens[587]. Based on the evidence that 80% of sgRNAs were successfully validated in the Rituximab arm, the 92 hits with FD >5 from the Rituximab screen were compared to gene sets in MSigDB, specifically to Hallmark and KEGG pathway gene sets and several pathways were found to be enriched with significant FDR values. Of particular note, TLR and MAP Kinase signalling were significantly enriched in the Rituximab arm. Of note, when the hits in the chemotherapy only arm were compared to MsigDB gene sets a different set of resistance mechanisms was identified with PI3K signalling the dominant enriched pathway. Due to the fact that Rituximab is acting without immune components in this *in vitro* system, it could be viewed as a direct chemotherapeutic, but despite this,

it appears to induce a different resistant phenotype to the other three combination chemotherapeutics.

An obvious limitation and cause for criticism in overexpression screens is the translational significance of the results. Indeed, ectopic alteration of the expression of a gene may induce a resistance phenotype in cell lines but if that gene is never altered in primary tumours in patients, then that potential resistance mechanism carries little relevance and benefit in the clinic. The agreement of some of our hits with well-known genes involved in BL and chemotherapy resistance such as PIM1 and ABC transporters was an encouraging observation as was the fact that components of the TLR and MAPK pathway were identified as being differentially expressed in pBL[62]. Furthermore, the endogenous expression of all genes at the validation stage was validated in the five PDX samples. These tumours were of course treatment-naïve but the fact that they are actively expressed in pBL samples supports the idea that they may be capable of driving a response to chemotherapy upon treatment. No clinical study of expression profile for treated pBL tumours was available to investigate this further, although, analysis of the 'relapsed' tumours in our study in chapter 6 highlighted that PIM1 was upregulated in all 5 PDXs, and the FDR value (0.061) fell just outside the stringent 0.05 threshold set for the study. Finally, Kaplan-Meier OS analysis based on gene expression levels of presentation samples was available for some of the SAM screen hits and although study numbers are relatively small and hence statistical significance is not always met, there is a trend towards worse prognosis associated with some of the identified hits in this screen. In particular, IRAK1 is one of the genes that gives the most significant association with reduced OS.

Overall, this SAM overexpression screen gave interesting insights into potential mechanisms of resistance in pBL but also highlighted some caveats and limitations associated with this form of prospective study. In particular, for our screen, the use of multi-agent chemotherapy proved a challenge. As long as a gene confers a survival advantage against one agent, it will give the cells some increased survival advantage in the screen setting and it will not be clear which agent/s it confers resistance to. So although multi-agent drug treatment is more clinically relevant, it requires laborious

validation with individual drugs. In this screen, there were no significant hits present in the chemotherapy + Rituximab arm only that might signify synergy between these drug treatments, rather the chemotherapy plus Rituximab arm was very similar to the chemotherapy alone arm. This may be due to the fact that the resistance mechanisms for both the three chemotherapeutics and Rituximab *in vitro* are separate but it is also possible that the drug concentrations were not optimal and hence the correct concentrations to use is a further complexity to be considered in the design of this type of screen. An alternative method could involve sequential treatment of SAM-ready cells with different drugs to ensure each individual drug is able to apply a strong selective pressure. However, as this screen was carried out over 13 days to ensure sufficient time for selection to occur this might pose an issue in sequential studies as the SAM activity may reduce over time. However, this was provisionally investigated in our lab and the SAM activity remained stable after four weeks. Alternatively, in order to maintain the use of clinically relevant combination immuno-chemotherapy a possible future step would be to perform a secondary screen with a library designed to contain just several hundred of the hits identified in this original screen with an increased number of sgRNAs per gene. This would minimise the laborious amount of scale up required for a large genome-wide library and importantly would enable further validation of hits across multiple cell lines and potentially at different concentrations of combination therapies. Indeed, recently it was suggested that this strategy should be adapted in screening experiments - perform a primary screen with a small number of sgRNAs per gene at genome-scale, use a relatively relaxed cut-off for hit selection and then perform a secondary screen on hundreds of primary screen hits using additional sgRNAs per gene[588]. As our screen was carried out in one cell line only it would be interesting to validate these hits in other cell lines before investigation in PDXs.

Other main concerns surrounding an overexpression screen involve the non-clinically relevant level of ectopic expression that can be induced. Indeed, in our original validation of 15 genes the fold change in expression of some genes was up to 3000 which is not a fold expression increase likely in a patient tumour. However, more relevant expression levels were seen in the majority of other sgRNA transduced cells. Despite the acknowledgments of the limitations and caveats that are associated with

overexpression screens, they can be a valuable tool to identify putative resistance mechanisms. Particularly in a cancer such a pBL, where relapse is detrimental and poorly understood, the ability to prospectively identify potential resistance mechanisms that can be further investigated, for example in *in vivo* models, is a valuable pre-clinical tool.

Chapter 8. Final discussion and future directions

8.1 Final discussion

Despite the exceptional improvements in curative outcome in pB-NHL over the last 3 decades salvaging of patients with relapse or refractory disease has not seen a similar improvement. This remains an inordinate challenge and there is an unmet need for improved therapeutic strategies. Data presented here has a broad scope with several different lines of investigation, all with the aim of identifying potential mechanisms of resistance in paediatric BL. Having discussed these data in the context of the literature at the end of each chapter, this chapter will combine the key findings and in particular address the therapeutic implications these discoveries may have.

8.1.1 Patient Derived Xenografts can be established for paediatric BL

The aim of chapter 3 was to establish a PDX resource to provide translationally relevant models for the study of heterogeneity and resistance mechanisms in pBL. PDXs have been shown to maintain the properties and heterogeneity of original patient samples and have much greater predictive power in translational research than cell lines[383]. PDXs were successfully established from pleural effusion, peripheral blood and bone marrow from patients with high-risk disease and significantly, in samples where detailed analysis of initial biopsy was possible, minute populations of cells were successfully maintained after passage in the mouse. Furthermore, serial passage of PDXs did not appear to impact the characteristics of the tumours with maintenance of histology, MYC translocation, gene expression and cell surface phonotype through multiple passages, highlighting the excellent clinical relevance of these models. To our knowledge this is the largest well-characterised paediatric BL PDX resource and offers excellent potential for pre-clinical research.

8.1.2 Heterogeneous populations exist in paediatric BL

The heterogenous nature of cancer is widely accepted at both the inter- and intra-tumour level and is considered one of the greatest challenges of cancer therapy. Despite the clinical significance of tumour heterogeneity, it has not been studied in paediatric

BL. Our well-characterised PDX models offer an excellent resource to investigate the different populations within each tumour and determine their clinical significance. Of particular note, the presence of cell surface proteins associated with stemness were identified in all PDXs including ABCG2 and CD34. Furthermore, PDX4 had a high proportion of CD90 positive cells, a HSC marker not usually present on mature B cells. Sorting of populations based on cell surface protein expression identified the varying levels of tumourigenic potential within PDXs, which is particularly striking in BL, a tumour with an extremely rapid turnover rate. For example, the CD9 negative population had significantly increased tumourigenic potential compared to the CD9 positive population across 4 PDXs, while other populations such as $CD44^+$ and $CD24^+$ cells also possessed increased tumourigenic potential in different PDXs. Despite the advantage conferred by some surface phenotypes, all PDXs maintained a constant population equilibrium with no outgrowth or loss of any population over 4 passages in mice. This highlights the potential importance of maintaining a tightly controlled equilibrium within the tumour. The intriguing level of heterogeneity was not just present at an intra-tumour level but also at an inter-tumour level with a distinct set of surface proteins expressed across the 5 PDXs. This heterogeneity is likely to be clinically relevant, indeed despite both PDXs being derived from pleural effusion of stage III BL, PDX4 was shown to be significantly more resistant to chemotherapy than PDX1. The heterogeneity was also evident at the molecular level and PCA of gene expression data in chapter 6 highlighted inter-patient heterogeneity and quite remarkably PDXs that were more closely related showed upregulation of common pathways following chemotherapy that was not observed in the other two PDXs.

8.1.3 A population with stemness exists in paediatric BL

It was not possible, based on surface protein expression alone, to identify a population with unique tumour propagating potential in the PDX models although a population with functional stemness was identified utilising the SP assay. Of particular importance, a SP was identified in the original biopsy sample from which PDX5 was derived, further validating the PDX model and highlighting that the SP is not an artefact of engraftment. Other biopsy samples were insufficient in size to carry out SP analysis, although

considering our emerging data showing the potential clinical relevance of this stem-like population, identification of an SP in biopsy samples received in the future will be a priority. The SP had significantly increased tumour initiating potential compared to MP cells *in vivo* and also gave rise to fully heterogenous tumours consisting of SP and MP cells. Further stemness features associated with the SP fraction include increased expression of ABC transporters and genes associated with anti-apoptotic programmes. Intriguingly, although all five PDXs showed upregulation of ABC transporters, there was heterogeneity with respect to other differentially expressed genes. For example NANOG expression was increased in the SP of only PDX1 while OCT4 was upregulated in the SP of two different PDXs and SOX2 expression was higher in the SP of three PDXs. Anti-apoptotic genes were differentially expressed in all five PDXs although the specific genes also varied for each SP. These data highlight that the SP phenotype may differ between tumours but they all converge on a stemness programme. The concept of a CSC in a mature B cell malignancy is not intuitive and the origin of this population must be considered. Significant evidence exists verifying the GC origin of the MYC translocation in BL. Gene expression profiling studies of BL samples have shown a strong relationship of BL cells with GC cells while DNA sequencing analysis showed that t(8;14) breakpoints are either in the switch regions or in the JH region, implying the involvement of VDJ recombination or SHM and immunoglobulin CSR[589,590]. Given the fact that the MYC translocation occurs in a GC B cell, a possible explanation for a CSC in BL is the acquisition of a mutation in a HSC or early progenitor followed by full malignant transformation in the GC. Indeed, transmission of premalignant lymphoid cells from 1 individual to another has been reported in at least 7 studies of mature lymphoid malignancies including CLL, FL and MCL although no evidence of this exists for BL or DLBCL[283–285]. An alternative explanation is the possibility that a cell harbouring the MYC translocation may undergo genetic or epigenetic reprogramming, conferring a stemness program on the background of a mature B cell which is the more likely scenario in light of the fact that GC cells can possess self-renewal potential that allows rapid lymphocyte division. The fact that the SP cells identified in this study are highly proliferative suggests that this population is more likely to arise from reprogramming of a mature, proliferative B cell rather than a quiescent progenitor. The growing body of evidence highlighting the

potential for the microenvironment experienced by tumour cells in lymph nodes and bone marrow to mediate stemness pathways as well as the exceptional plasticity demonstrated by B cells provide support for a this reprogramming hypothesis[591]. The exact cell of origin of the SP in BL will be an exciting project to pursue.

Regardless of the exact origin, the SP in our study has several characteristics of a CSC including significantly increased tumourigenic potential compared to MP cells as well as the ability to produce fully heterogenous tumours composed of both SP and MP cells. They also express high levels of ABC transporter proteins and anti-apoptotic programmes. However, the MP cells are not completely devoid of tumourigenic potential and it was not possible to prospectively isolate the exact population by flow cytometry using cell surface proteins, an important characteristic of the CSC according to Dick and colleagues[202]. Hence, although we cannot assign the term CSC to the SP, this population does undoubtedly represent a heterogenous compartment with increased stemness in BL. Koch *et al.* in their study of SP cells in DLBCL suggest the existence of a self-maintained intra-clonal equilibrium of tumour cell populations[230]. Indeed, our study of heterogenous populations in BL suggest a similar tightly controlled equilibrium. All tumour populations are maintained at a constant level regardless of the increased tumourigenic potential of the SP and other populations identified by surface protein expression levels. Furthermore, *in vitro* the minority SP was shown to significantly influence the bulk cell population's growth, via soluble factors, highlighting the importance of maintenance of a very minor population for overall tumour growth dynamics. Significantly, we were able to identify differentially produced cytokines between the SP and MP, suggesting this is one way the equilibrium between SP and MP cells is maintained. Interestingly, although the exchange of soluble factors between the SP and MP is essential for MP proliferative capabilities during early culture, following several weeks a SP is present in MP-derived cell lines, further supporting the idea that there may be plasticity and interconversion of MP to SP cells under certain conditions. Irrespective of whether a distinct hierarchical organisation exists in BL or not, the data from this thesis highlight that a tightly controlled sub-clonal architecture is present in BL and the disease is not as homogenous as assumed.

8.1.4 Inhibition of PI3K/mTOR signalling may be a therapeutic strategy for targeting the SP

An aim of this thesis was to determine if heterogenous populations within BL could have clinical significance. SP cells within cell lines and PDXs have increased propagating potential with *in vitro* evidence highlighting that soluble factors produced by the SP are important for MP growth. As well as this, SP cells have significantly increased resistance to chemotherapy compared to the normal whole cell population and evidence that this population may be important in driving resistance to chemotherapy is provided by the enrichment of this population in the 'mimicked' relapsed PDXs. Given the significant data pointing to an essential role for the SP in tumour dynamics and response to chemotherapy, the ability to target this population could have important clinical implications. *In vitro* analysis highlighted that inhibition of PI3K/mTOR signalling diminished the SP phenotype by downregulating expression of ABCG2 and ABCB1 and significantly it had a synergistic interaction with doxorubicin and methotrexate in the treatment of the Raji cell line. PI3K signalling was also enriched in three PDX tumours following the induction of relapse. This increase cannot be attributed specifically to the SP fraction although it provides further support for the important role of this pathway in mediating a resistance phenotype in BL. Furthermore, in the CRISPR screen, chemotherapy treatment resulted in enrichment of hits associated with PI3K/AKT/mTOR signalling. With all evidence pointing to the important role of this pathway in the SP and drug resistance phenotypes it will be important to further investigate this target in the PDX models. The PI3K inhibitor Idelalisib has shown an objective response in phase II data evaluating adults with relapsed indolent mature B-NHL, as well as significantly improved outcomes in combination with Rituximab in adults with relapsed CLL[112]. Interestingly, a phase II trial of Idelalisib in children and adolescents with relapsed or refractory DLBCL or mediastinal B-cell lymphoma has just recently started (NCT03349346) but does not include BL. More pre-clinical data such as ours may lead to a trial for BL in the future.

8.1.5 Other potential therapeutic strategies in resistant disease were identified

A number of lines of investigation in this thesis highlight the possibility that ABC transporters are important mediators of the response to chemotherapy in BL. Not only does ABCG2 and ABCB1 define the SP but ABCB1 was significantly upregulated in all relapsed PDXs and ectopic expression of it, and ABCC1 in the overexpression screen also strongly mediated resistance to chemotherapy. Several chemotherapy drugs were not included in our studies and indeed these drugs, particularly prednisolone, should be studied in the context of the SP and ABC transporter expression in future experiments. Despite this, these significant data suggest that ABC transporter targeting should be investigated in BL. Initial clinical trials involving three generations of inhibitors of ABC transporters were not successful, although a new era of research in this area has presented some strategies with increased promise[592]. Examples of improved approaches include inactivation of ABC transporter associated genes by microRNA and RNA interference as well as delivery of these microRNAs or other anti-cancer drugs directly to resistant cells by nanoparticles[155]. Monoclonal antibodies targeting the encoded P-glycoprotein also exist and one is currently in clinical trials for leukaemia[593]. Enrichment of the B cell receptor pathway in a subset of PDXs after relapse is a further interesting insight as it provides evidence of the translational relevance of our PDX relapse model, particularly in light of the establishment of an Ibrutinib clinical trial in relapse and refractory paediatric B-NHL (NCT02703272). Interestingly, Ibrutinib was shown to sensitize primary resistant cells and cell lines to chemotherapeutics by inhibiting P-glycoprotein ATPase activity and hence further investigation of the impact of this drug on the SP fraction in our PDXs is warranted[594].

8.1.6 A CRISPR-Cas9 overexpression screen prospectively identified putative drivers of resistance to both chemotherapeutics and Rituximab

A genome-wide CRISPR-Cas9 overexpression screen was carried out to provide a broad overview of putative resistance mechanisms to both chemotherapy and Rituximab that could then be further validated in our PDX model. As well as the very significant resistance conferred by ectopic expression of ABC transporters and enrichment of PI3K signalling components, other potential targets such as PIM1 and PEG10 were drivers of

resistance to the chemotherapeutic agents. Of note, PIM1 was upregulated in the SP of four out of five of our PDXs and should be investigated further. In the Rituximab arm, a number of interesting resistance mechanisms were also elucidated that warrant further investigation. Rituximab has now been integrated into frontline therapy for pBL and although overall survival rate is excellent, some patients still relapse and therefore it is necessary to identify novel strategies that may mediate resistance to this drug. We acknowledge that without the presence of a functioning immune system, the data from this model are preliminary and will need to be investigated in a more clinically relevant model. Regardless, IRAK1 expression was shown to strongly induce resistance to Rituximab in the *in vitro* screen setting. This gene was recently reported to be upregulated for the first time in a large study of untreated paediatric BL[62]. It will be interesting to investigate if this gene, or indeed the identified MAPK and TLR signalling pathways, drive resistance to Rituximab in our PDXs, for example, in a humanised mouse model.

8.2 Future work

8.2.1 Further characterisation of the PDX models

Given the excellent translational potential of our PDX resource, further methods will be undertaken to fully validate this model. Whole exome sequencing will be carried out comparing DNA from the original patient biopsy samples to DNA extracted from tumours at various passages in mice. This will allow us to investigate the genetic stability of these models and elucidate if any mutations occur as a result of engraftment and adaption to *in vivo* murine growth while also determining if populations are being selected or lost due to engraftment.

In order to incorporate Rituximab into our future *in vivo* experiments we will consider the use of humanised NSG mice. For example, huCD34^{+} NSG mice which reconstitute mice with the major cell types of the human immune system would allow for investigation of resistance mechanisms to Rituximab in the context of ADCC and CMC and would allow us to validate the putative resistance mechanisms to Rituximab

identified in chapter 7 such as MAPK and TLR signalling in a more clinically relevant model.

8.2.2 Investigation of targeted therapies in relapsed PDXs

Given the successful establishment of an *in vivo* model of relapse we will further develop this project to determine effective therapeutics for relapse disease. Three chemotherapeutic agents (doxorubicin, methotrexate and vincristine) were used in this project and the effective concentrations determined. Future studies may incorporate further drugs including prednisolone and Rituximab, the latter of course with the caveats of requiring a functional immune system. Following the emergence of relapse of the PDX tumours, the efficacy of different drugs will be investigated including Ibrutinib and Idelalisib as well as inhibitors of anti-apoptotic proteins such as PIM1 kinase and BCL2, particularly given the recurring evidence in this thesis that anti-apoptotic programmes are important in resistance in BL. This will provide valuable pre-clinical data that is currently lacking in the field of relapse in pBL. Given the striking increased resistance of the PDX SP to chemotherapy, Ibrutinib, Dactolisib, Idelalisib and PIM1 kinase inhibitors will also be compared to chemotherapy to determine the most effective mechanism of targeting this population. Based on our *in vitro* data we would expect Dactolisib to be effective at targeting the SP and studies involving Idelalisib will be informative to determine whether PI3K or mTOR signalling is more important in mediating the SP phenotype through ABC transporter regulation. Indeed, as it is unlikely that targeted therapies will be utilised in paediatric BL as single agents the effect of these drugs combined with chemotherapy will be investigated and the potential for synergistic combinations will be determined, as was noted for chemotherapy plus Dactolisib *in vitro.* We have also recently acquired a PDX model derived from an *in situ* relapsed tumour through a collaboration which, when established in our lab, will offer an excellent resource to compare with our PDX models of relapse. In particular, we will examine this tumour for the presence of a SP. Of note, a SP is present in all our PDXs derived from tumours at primary presentation although clinically, only a small number of BL patients relapse. It would therefore be particularly informative to compare the SP

of a *bone fide* relapsed sample with the SP of patients who did not relapse, for example by RNA-seq, to determine if any unique characteristics can be identified.

8.2.3 Investigation of the clonal dynamics of pBL

Given the fact that we have highlighted the potential clinical significance of the stemness population in pBL but have not yet determined the exact origin of this population, clonal dynamics of BL will be investigated. We have recently acquired FFPE material of a presentation and relapse pBL sample along with matched healthy tissue. High coverage multi-region whole exome or genome sequencing will be carried out on these paired samples. Phylogenetic trees will be inferred to depict the evolutionary history of the tumour, and therefore unveil the temporal and spatial occurrence of mutations, and the origins of a population that leads to relapse.

In vivo the clonal dynamics within PDXs will be investigated to determine if the increased number of SP cells identified in our relapsed PDX models are directly related to the SP in the untreated tumour or if therapy resulted in acquisition of *de novo* SP cells. Lineage tracing or fate mapping will be utilised to identify all the progeny of the SP within the tumour cell following chemotherapy. Traditional methods to trace cell lineage such as vital dyes, fluorescent proteins and DNA tags may be sufficient to monitor the clonal dynamics of the SP versus the bulk tumour population with and without chemotherapy. More modern and sophisticated methods that harness the power of CRISPR cas9 such as genome editing of synthetic target arrays for lineage tracing (GESTALT) will also be considered[594]. This method uses genome editing to generate a combinatorial diversity of mutations that accumulate over many cell divisions so that each progeny possess a unique mark and is not identical, enabling more detailed analysis of cellular relationships. Determination of the best method to use will likely involve the outcome of efforts to grow our PDXs *ex vivo* (see below).

8.2.4 Identification of SP-derived soluble factors that mediate MP growth

The identification of differentially expressed cytokines in the SP of three cell lines is an encouraging indicator that we may be able to elucidate the mechanisms of cross-talk

between SP and MP cells. Future experiments will involve the use of further cytokine panels to produce a more exhaustive overview of the cytokine profiles of the SP and MP cells in each cell line. This may lead to the identification of cytokines that are increased in the SP of all cell lines. We will devise a matrix approach to supplement MP cells with the cytokines that were identified as significantly increased in the SP in an attempt to recover MP growth. The data produced so far was produced from studies of cell lines hence we will endeavour to measure the cytokine production by MP and SP cells from PDXs *ex vivo*. It has thus far not been possible to grow the PDXs *in vitro* although we have recently acquired a feeder cell layer that has proven effective at supporting the growth of the tumours *ex vivo,* based on the expression of CD40L and IL21 production. After further validation of this resource, particularly verification that it does not induce phenotypic or genetic changes of the PDX-derived cells, we will culture SP and whole population cells *ex vivo* and measure the production of cytokines in both populations, starting with cytokines identified as potential mediators of MP growth in the cell line studies. As well as informing us about the biology of these tumour cells, identification of cytokines involved in driving bulk tumour growth would present potential therapeutic opportunities.

Overall, this study has provided an interesting insight into the heterogeneity within pBL and the potential clinical implications of this heterogeneity. Furthermore, it has identified mechanisms of resistance to therapies and has provided significant groundwork for future studies. Completion of the above lines of investigation will likely result in significant improvements to our understanding of the biology of relapse in pBL.

Chapter 9. Bibliography

1. Swerdlow, S., Campo, E., Pileri, S., Harris, N. & Stein, H. The 2016 revision of the World Health Organisation classification of lymphoid neoplasms. *Blood* **127,** 2375–2390 (2016).
2. Https://www.cancerresearchuk.org/health-professional/cancer-statistics/. CRUK Non-Hodgkin Lymphoma incidence statistics.
3. Sandlund, J. T., Downing, J. R. & Crist, W. M. Non-Hodgkin's Lymphoma in Childhood. *N. Engl. J. Med.* **334,** 1238–1248 (1996).
4. Burkhardt, B. *et al.* The impact of age and gender on biology, clinical features and treatment outcome of non-Hodgkin lymphoma in childhood and adolescence. *Br. J. Haematol.* **131,** 39–49 (2005).
5. Lange, J. & Burkhardt, B. Treatment of Adolescents with Aggressive B-Cell Malignancies: The Pediatric Experience. *Curr. Hematol. Malig. Rep.* **8,** 226–235 (2013).
6. Hochberg, J., Waxman, I. M., Kelly, K. M., Morris, E. & Cairo, M. S. Adolescent non-Hodgkin lymphoma and Hodgkin lymphoma: state of the science. *Br. J. Haematol.* **144,** 24–40 (2009).
7. Pieper, K., Grimbacher, B. & Eibel, H. B-cell biology and development. *J. Allergy Clin. Immunol.* **131,** 959–971 (2013).
8. Morton, L. M. *et al.* Lymphoma incidence patterns by WHO subtype in the United States, 1992-2001. *Blood* **107,** 265–276 (2006).
9. Sant, M. *et al.* Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood* **116,** 3724–3734 (2010).
10. Küppers, R., Klein, U., Hansmann, M.-L. & Rajewsky, K. Cellular Origin of Human B-Cell Lymphomas. *N. Engl. J. Med.* **341,** 1520–1529 (1999).
11. Eibel, H., Kraus, H., Sic, H., Kienzler, A.-K. & Rizzi, M. B cell Biology: An Overview. *Curr. Allergy Asthma Rep.* **14,** 434 (2014).
12. Victora, G. D. & Nussenzweig, M. C. Germinal Centers. *Annu. Rev. Immunol.* **30,** 429–457 (2012).
13. Basso, K. & Dalla-Favera, R. Germinal centres and B cell lymphomagenesis. *Nat. Rev. Immunol.* **15,** 172–184 (2015).
14. Victora, G. D. *et al.* Identification of human germinal center light and dark zone cells and their relationship to human B-cell lymphomas. *Blood* **120,** 2240–2248 (2012).
15. Küppers, R. Mechanisms of B-cell lymphoma pathogenesis. *Nat. Rev. Cancer* **5,** 251–262 (2005).
16. Richter, J. *et al.* Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nat. Genet.* **44,** 1316–1320 (2012).
17. Schmitz, R. *et al.* Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature* **490,** 116–120 (2012).

18. Zhang, S. *et al.* Identification and Characterization of Ovarian Cancer-Initiating Cells from Primary Human Tumors. *Cancer Res.* **68,** 4311–4320 (2008).

19. Houghton, P. J. *et al.* The pediatric preclinical testing program: Description of models and early testing results. *Pediatr. Blood Cancer* **49,** 928–940 (2007).

20. Küppers, R. & Dalla-Favera, R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene* **20,** 5580–5594 (2001).

21. El-Mallawany, N. K. & Cairo, M. S. Advances in the diagnosis and treatment of childhood and adolescent B-cell non-Hodgkin lymphoma. *Clin. Adv. Hematol. Oncol.* **13,** 113–23 (2015).

22. Burkitt, D. A sarcoma involving the jaws in african children. *Br. J. Surg.* **46,** 218–223 (1958).

23. EPSTEIN, M. A., ACHONG, B. G. & BARR, Y. M. VIRUS PARTICLES IN CULTURED LYMPHOBLASTS FROM BURKITT'S LYMPHOMA. *Lancet (London, England)* **1,** 702–3 (1964).

24. Zech, L., Haglund, U., Nilsson, K. & Klein, G. Characteristic chromosomal abnormalities in biopsies and lymphoid-cell lines from patients with Burkitt and non-Burkitt lymphomas. *Int. J. cancer* **17,** 47–56 (1976).

25. Manolov, G. & Manolova, Y. Marker band in one chromosome 14 from Burkitt lymphomas. *Nature* **237,** 33–4 (1972).

26. Dalla-Favera, R. *et al.* Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc. Natl. Acad. Sci. U. S. A.* **79,** 7824–7 (1982).

27. Lenze, D. *et al.* The different epidemiologic subtypes of Burkitt lymphoma share a homogenous micro RNA profile distinct from diffuse large B-cell lymphoma. *Leukemia* **25,** 1869–1876 (2011).

28. Piccaluga, P. P. *et al.* Gene expression analysis uncovers similarity and differences among Burkitt lymphoma subtypes. *Blood* **117,** 3596–608 (2011).

29. Amato, T. *et al.* Clonality Analysis of Immunoglobulin Gene Rearrangement by Next-Generation Sequencing in Endemic Burkitt Lymphoma Suggests Antigen Drive Activation of BCR as Opposed to Sporadic Burkitt Lymphoma. *Am. J. Clin. Pathol.* **145,** 116–127 (2016).

30. Abate, F. *et al.* Distinct Viral and Mutational Spectrum of Endemic Burkitt Lymphoma. *PLOS Pathog.* **11,** e1005158 (2015).

31. Dalldorf, G., Linsell, C. A., Barnhart, F. E. & Martyn, R. An Epidemiologic Approach to the Lymphomas of African Children and Burkitt's Sarcoma of the Jaws. *Perspect. Biol. Med.* **7,** 435–449 (1964).

32. Morrow, R. H., Kisuule, A., Pike, M. C. & Smith, P. G. Burkitt's lymphoma in the Mengo Districts of Uganda: epidemiologic features and their relationship to malaria. *J. Natl. Cancer Inst.* **56,** 479–83 (1976).

33. Illingworth, J. *et al.* Chronic Exposure to Plasmodium falciparum Is Associated with Phenotypic Evidence of B and T Cell Exhaustion. *J. Immunol.* **190,** 1038–1047 (2013).

34. Donati, D. *et al.* Clearance of Circulating Epstein-Barr Virus DNA in Children with Acute Malaria after Antimalaria Treatment. *J. Infect. Dis.* **193,** 971–977 (2006).

35. Chêne, A. *et al.* A Molecular Link between Malaria and Epstein–Barr Virus Reactivation. *PLoS Pathog.* **3,** e80 (2007).

36. Mutalima, N. *et al.* Associations between Burkitt Lymphoma among Children in Malawi and Infection with HIV, EBV and Malaria: Results from a Case-Control Study. *PLoS One* **3,** e2505 (2008).

37. Bornkamm, G. W. Epstein-Barr virus and the pathogenesis of Burkitt's lymphoma: More questions than answers. *Int. J. Cancer* **124,** 1745–1755 (2009).

38. Thorley-Lawson, D. A. Epstein-Barr virus: exploiting the immune system. *Nat. Rev. Immunol.* **1,** 75–82 (2001).

39. Schmitz, R., Ceribelli, M., Pittaluga, S., Wright, G. & Staudt, L. M. Oncogenic Mechanisms in Burkitt Lymphoma. *Cold Spring Harb. Perspect. Med.* **4,** a014282–a014282 (2014).

40. Boerma, E. G. *et al.* Gender and age-related differences in Burkitt lymphoma – epidemiological and clinical data from The Netherlands. *Eur. J. Cancer* **40,** 2781–2787 (2004).

41. Szczepanowski, M. *et al.* Cell-of-origin classification by gene expression and *MYC* -rearrangements in diffuse large B-cell lymphoma of children and adolescents. *Br. J. Haematol.* **179,** 116–119 (2017).

42. Schulz, T. F., Boshoff, C. H. & Weiss, R. A. HIV infection and neoplasia. *Lancet* **348,** 587–591 (1996).

43. Ferry, J. A. Burkitt's Lymphoma: Clinicopathologic Features and Differential Diagnosis. *Oncologist* **11,** 375–383 (2006).

44. Molyneux, E. M. *et al.* Burkitt's lymphoma. *Lancet* **379,** 1234–1244 (2012).

45. Carbone, A. Emerging pathways in the development of AIDS-related lymphomas. *Lancet. Oncol.* **4,** 22–9 (2003).

46. Fisher, S. G. & Fisher, R. I. The epidemiology of non-Hodgkin's lymphoma. *Oncogene* **23,** 6524–6534 (2004).

47. Hecht, J. L. & Aster, J. C. Molecular Biology of Burkitt's Lymphoma. *J. Clin. Oncol.* **18,** 3707–3721 (2000).

48. ar-Rushdi, A. *et al.* Differential expression of the translocated and the untranslocated c-myc oncogene in Burkitt lymphoma. *Science* **222,** 390–3 (1983).

49. Fahrlandertt, P. D. *et al. Activation of the c-myc oncogene by the immunoglobulin heavy-chain gene enhancer after multiple switch region-mediated chromosome rearrangements in a murine plasmacytoma (chromosome translocation)*. *Proc. Nati. Acad. Sci. USA* **82,** (1985).

50. Janz, S., Potter, M. & Rabkin, C. S. Lymphoma- and leukemia-associated chromosomal translocations in healthy individuals. *Genes, Chromosom. Cancer* **36,** 211–223 (2003).

51. Adams, J. M. *et al.* The c-myc oncogene driven by immunoglobulin enhancers

induces lymphoid malignancy in transgenic mice. *Nature* **318,** 533–538 (1985).

52. Love, C. *et al.* The genetic landscape of mutations in Burkitt lymphoma. *Nat. Genet.* **44,** 1321–1325 (2012).

53. Hemann, M. T. *et al.* Evasion of the p53 tumour surveillance network by tumour-derived MYC mutants. *Nature* **436,** 807–811 (2005).

54. Evan, G. I. *et al.* Induction of apoptosis in fibroblasts by c-myc protein. *Cell* **69,** 119–28 (1992).

55. Meyer, N., Kim, S. S. & Penn, L. Z. The Oscar-worthy role of Myc in apoptosis. *Semin. Cancer Biol.* **16,** 275–287 (2006).

56. Rohde, M. *et al.* Relevance of ID3-TCF3-CCND3 pathway mutations in pediatric aggressive B-cell lymphoma treated according to the non-Hodgkin Lymphoma Berlin-Frankfurt-Münster protocols. *Haematologica* **102,** 1091–1098 (2017).

57. Young, R. M. & Staudt, L. M. Targeting pathological B cell receptor signalling in lymphoid malignancies. *Nat. Rev. Drug Discov.* **12,** 229–243 (2013).

58. Dave, S. S. *et al.* Molecular Diagnosis of Burkitt's Lymphoma. *N. Engl. J. Med.* **354,** 2431–2442 (2006).

59. Sander, S. *et al.* Synergy between PI3K Signaling and MYC in Burkitt Lymphomagenesis. *Cancer Cell* **22,** 167–179 (2012).

60. Xiao, C. *et al.* Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes. *Nat. Immunol.* **9,** 405–414 (2008).

61. Olive, V. *et al.* miR-19 is a key oncogenic component of mir-17-92. *Genes Dev.* **23,** 2839–2849 (2009).

62. Lee, S. *et al.* Comparative genomic expression signatures of signal transduction pathways and targets in paediatric Burkitt lymphoma: a Children's Oncology Group report. *Br. J. Haematol.* **177,** 601–611 (2017).

63. Kim, H. *et al.* Clinical outcome of relapsed or refractory burkitt lymphoma and mature B-cell lymphoblastic leukemia in children and adolescents. *Cancer Res. Treat.* **46,** 358–65 (2014).

64. Poirel, H. A. *et al.* Specific cytogenetic abnormalities are associated with a significantly inferior outcome in children and adolescents with mature B-cell non-Hodgkin's lymphoma: results of the FAB/LMB 96 international study. *Leukemia* **23,** 323–331 (2009).

65. Nelson, M. *et al.* An increased frequency of 13q deletions detected by fluorescence *in situ* hybridization and its impact on survival in children and adolescents with Burkitt lymphoma: results from the Children's Oncology Group study CCG-5961. *Br. J. Haematol.* **148,** 600–610 (2010).

66. Salaverria, I. *et al.* Chromosomal alterations detected by comparative genomic hybridization in subgroups of gene expression-defined Burkitt's lymphoma. *Haematologica* **93,** 1327–34 (2008).

67. Toujani, S. *et al.* High Resolution Genome-Wide Analysis of Chromosomal

Alterations in Burkitt's Lymphoma. *PLoS One* **4,** e7089 (2009).

68. Scholtysik, R. *et al.* Detection of genomic aberrations in molecularly defined Burkitt's lymphoma by array-based, high resolution, single nucleotide polymorphism analysis. *Haematologica* **95,** 2047–2055 (2010).

69. Schiffman, J. D. *et al.* Genome wide copy number analysis of paediatric Burkitt lymphoma using formalin-fixed tissues reveals a subset with gain of chromosome 13q and corresponding miRNA over expression. *Br. J. Haematol.* **155,** 477–86 (2011).

70. He, L. *et al.* A microRNA polycistron as a potential human oncogene. *Nature* **435,** 828–833 (2005).

71. Tagawa, H., Karube, K., Tsuzuki, S., Ohshima, K. & Seto, M. Synergistic action of the microRNA-17 polycistron and Myc in aggressive cancer development. *Cancer Sci.* **98,** 1482–1490 (2007).

72. O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V. & Mendell, J. T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435,** 839–843 (2005).

73. Diosdado, B. *et al.* MiR-17-92 cluster is associated with 13q gain and c-myc expression during colorectal adenoma to adenocarcinoma progression. *Br. J. Cancer* **101,** 707–714 (2009).

74. Chacon-Cortes, D., Smith, R. A., Lea, R. A., Youl, P. H. & Griffiths, L. R. Association of microRNA 17–92 cluster host gene (MIR17HG) polymorphisms with breast cancer. *Tumor Biol.* **36,** 5369–5376 (2015).

75. Hammond, S. M. MicroRNAs as oncogenes. *Curr. Opin. Genet. Dev.* **16,** 4–9 (2006).

76. Campo, E. *et al.* The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. *Blood* **117,** 5019–32 (2011).

77. Gifford, G. K., Gill, A. J. & Stevenson, W. S. Molecular subtyping of diffuse large B-cell lymphoma: update on biology, diagnosis and emerging platforms for practising pathologists. *Pathology* **48,** 5–16 (2016).

78. Hans, C. P. *et al.* Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* **103,** 275–282 (2004).

79. Miles, R. R. *et al.* Pediatric diffuse large B-cell lymphoma demonstrates a high proliferation index, frequent c-Myc protein expression, and a high incidence of germinal center subtype: Report of the French-American-British (FAB) international study group. *Pediatr. Blood Cancer* **51,** 369–374 (2008).

80. Sehn, L. H. & Gascoyne, R. D. Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity. *Blood* **125,** 22–32 (2015).

81. Pasqualucci, L. *et al.* Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet.* **43,** 830–7 (2011).

82. Zhang, J. *et al.* Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc. Natl. Acad. Sci.* **110,** 1398–1403 (2013).

83. Davis, R. E. *et al.* Chronic active B-cell-receptor signalling in diffuse large B-cell

lymphoma. *Nature* **463,** 88–92 (2010).

84. Iqbal, J. *et al.* BCL2 translocation defines a unique tumor subset within the germinal center B-cell-like diffuse large B-cell lymphoma. *Am. J. Pathol.* **165,** 159–66 (2004).

85. Lenz, G. *et al.* Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 13520–5 (2008).

86. Salaverria, I. *et al.* Translocations activating IRF4 identify a subtype of germinal center-derived B-cell lymphoma affecting predominantly children and young adults. *Blood* **118,** 139–47 (2011).

87. Klapper, W. *et al.* Molecular profiling of pediatric mature B-cell lymphoma treated in population-based prospective clinical trials. *Blood* **112,** 1374–1381 (2008).

88. Reiter, A. & Klapper, W. Recent advances in the understanding and management of diffuse large B-cell lymphoma in children. *Br. J. Haematol.* **142,** 329–347 (2008).

89. Cairo, M. S. *et al.* Results of a randomized international study of high risk central nervous system B-non-Hodgkin's lymphoma and B-acute lymphoblastic leukemia in children and adolescents. *Blood* **109,** 2736–43 (2006).

90. Patte, C. *et al.* Results of the randomized international FAB/LMB96 trial for intermediate risk B-cell non-Hodgkin's lymphoma in children and adolescents: it is possible to reduce treatment for the early responding patients. *Blood* **109,** 2773–80 (2006).

91. Gerrard, M. *et al.* Excellent survival following two courses of COPAD chemotherapy in children and adolescents with resected localized B-cell non-Hodgkin's lymphoma: results of the FAB/LMB 96 international study. *Br. J. Haematol.* **141,** 840–847 (2008).

92. Kliegman, R. & Nelson, W. E. (Waldo E. *Nelson textbook of pediatrics.* (Elsevier/Saunders, 2011).

93. Miles, R. R., Arnold, S. & Cairo, M. S. Risk factors and treatment of childhood and adolescent Burkitt lymphoma/leukaemia. *Br. J. Haematol.* **156,** 730–43 (2012).

94. Coiffier, B. *et al.* CHOP Chemotherapy plus Rituximab Compared with CHOP Alone in Elderly Patients with Diffuse Large-B-Cell Lymphoma. *N. Engl. J. Med.* **346,** 235–242 (2002).

95. Habermann, T. M. *et al.* Rituximab-CHOP Versus CHOP Alone or With Maintenance Rituximab in Older Patients With Diffuse Large B-Cell Lymphoma. *J. Clin. Oncol.* **24,** 3121–3127 (2006).

96. Pfreundschuh, M. *et al.* CHOP-like chemotherapy with or without rituximab in young patients with good-prognosis diffuse large-B-cell lymphoma: 6-year results of an open-label randomised study of the MabThera International Trial (MInT) Group. *Lancet Oncol.* **12,** 1013–1022 (2011).

97. Mounier, N. *et al.* Rituximab plus CHOP (R-CHOP) overcomes bcl-2--associated resistance to chemotherapy in elderly patients with diffuse large B-cell lymphoma (DLBCL). *Blood* **101,** 4279–84 (2003).

98. Winter, J. N. *et al.* Prognostic significance of Bcl-6 protein expression in DLBCL

treated with CHOP or R-CHOP: a prospective correlative study. *Blood* **107,** 4207–13 (2006).

99. Ribrag, V. *et al.* Rituximab and dose-dense chemotherapy for adults with Burkitt's lymphoma: a randomised, controlled, open-label, phase 3 trial. *Lancet (London, England)* **387,** 2402–11 (2016).

100. Goldman, S. *et al.* Rituximab and FAB/LMB 96 chemotherapy in children with Stage III/IV B-cell non-Hodgkin lymphoma: a Children's Oncology Group report. *Leukemia* **27,** 1174–7 (2013).

101. Goldman, S. *et al.* Rituximab with chemotherapy in children and adolescents with central nervous system and/or bone marrow-positive Burkitt lymphoma/leukaemia: a Children's Oncology Group Report. *Br. J. Haematol.* **167,** 394–401 (2014).

102. Minard-Colin, V. *et al.* Results of the randomized Intergroup trial Inter-B-NHL Ritux 2010 for children and adolescents with high-risk B-cell non-Hodgkin lymphoma (B-NHL) and mature acute leukemia (B-AL): Evaluation of rituximab (R) efficacy in addition to standard LMB chemotherapy (CT) regimen. *J. Clin. Oncol.* **34,** 10507–10507 (2016).

103. Osumi, T. *et al.* Improved Outcome of Relapsed or Refractory Pediatric B-Cell Non-Hodgkin Lymphoma in Japan. *Blood* **126,** (2015).

104. Herter, S. *et al.* Preclinical Activity of the Type II CD20 Antibody GA101 (Obinutuzumab) Compared with Rituximab and Ofatumumab In Vitro and in Xenograft Models. *Mol. Cancer Ther.* **12,** 2031–2042 (2013).

105. Younes, A. *et al.* Phase I Multidose-Escalation Study of the Anti-CD19 Maytansinoid Immunoconjugate SAR3419 Administered by Intravenous Infusion Every 3 Weeks to Patients With Relapsed/Refractory B-Cell Lymphoma. *J. Clin. Oncol.* **30,** 2776–2782 (2012).

106. Grupp, S. A. *et al.* Chimeric Antigen Receptor–Modified T Cells for Acute Lymphoid Leukemia. *N. Engl. J. Med.* **368,** 1509–1518 (2013).

107. Porter, D. L., Levine, B. L., Kalos, M., Bagg, A. & June, C. H. Chimeric Antigen Receptor–Modified T Cells in Chronic Lymphoid Leukemia. *N. Engl. J. Med.* **365,** 725–733 (2011).

108. Advani, R. H. *et al.* Bruton Tyrosine Kinase Inhibitor Ibrutinib (PCI-32765) Has Significant Activity in Patients With Relapsed/Refractory B-Cell Malignancies. *J. Clin. Oncol.* **31,** 88–94 (2013).

109. Byrd, J. C. *et al.* Targeting BTK with Ibrutinib in Relapsed Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **369,** 32–42 (2013).

110. Wang, M. L. *et al.* Targeting BTK with Ibrutinib in Relapsed or Refractory Mantle-Cell Lymphoma. *N. Engl. J. Med.* **369,** 507–516 (2013).

111. Wilson, W. H. *et al.* The Bruton's Tyrosine Kinase (BTK) Inhibitor, Ibrutinib (PCI-32765), Has Preferential Activity in the ABC Subtype of Relapsed/Refractory De Novo Diffuse Large B-Cell Lymphoma (DLBCL): Interim Results of a Multicenter, Open-Label, Phase 2 Study. *Blood* **120,** (2012).

112. Furman, R. R. *et al.* Idelalisib and Rituximab in Relapsed Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **370,** 997–1007 (2014).

113. Weiss, R. B. The anthracyclines: will we ever find a better doxorubicin? *Semin. Oncol.* **19,** 670–86 (1992).

114. Tewey, K. M., Rowe, T. C., Yang, L., Halligan, B. D. & Liu, L. F. Adriamycin-induced DNA damage mediated by mammalian DNA topoisomerase II. *Science* **226,** 466–8 (1984).

115. Germann, U. A. P-glycoprotein--a mediator of multidrug resistance in tumour cells. *Eur. J. Cancer* **32A,** 927–44 (1996).

116. Cole, S. P. *et al.* Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science* **258,** 1650–4 (1992).

117. Young, L. C., Campling, B. G., Cole, S. P., Deeley, R. G. & Gerlach, J. H. Multidrug resistance proteins MRP3, MRP1, and MRP2 in lung cancer: correlation of protein levels with drug response and messenger RNA levels. *Clin. Cancer Res.* **7,** 1798–804 (2001).

118. Burgess, D. J. *et al.* Topoisomerase levels determine chemotherapy response *in vitro* and *in vivo*. *Proc. Natl. Acad. Sci.* **105,** 9053–9058 (2008).

119. Gewirtz, D. A. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. *Biochem. Pharmacol.* **57,** 727–41 (1999).

120. Doroshow, J. H. Role of hydrogen peroxide and hydroxyl radical formation in the killing of Ehrlich tumor cells by anticancer quinones. *Proc. Natl. Acad. Sci. U. S. A.* **83,** 4514–8 (1986).

121. Naci, D. *et al.* α2β1 Integrin Promotes Chemoresistance against Doxorubicin in Cancer Cells through Extracellular Signal-regulated Kinase (ERK). *J. Biol. Chem.* **287,** 17065–17076 (2012).

122. Bhatti, M. *et al.* Pre-clinical activity of targeting the PI3K/Akt/mTOR pathway in Burkitt lymphoma. *Oncotarget* **9,** 21820–21830 (2018).

123. Chan, E. S. L. & Cronstein, B. N. *Mechanisms of action of methotrexate*. *Bulletin of the Hospital for Joint Diseases* **71,** (2013).

124. Laverdiere, C., Chiasson, S., Costea, I., Moghrabi, A. & Krajinovic, M. Polymorphism G80A in the reduced folate carrier gene and its relationship to methotrexate plasma levels and outcome of childhood acute lymphoblastic leukemia. *Blood* **100,** 3832–3834 (2002).

125. Drori, S., Jansen, G., Mauritz, R., Peters, G. J. & Assaraf, Y. G. Clustering of Mutations in the First Transmembrane Domain of the Human Reduced Folate Carrier in GW1843U89-resistant Leukemia Cells with Impaired Antifolate Transport and Augmented Folate Uptake. *J. Biol. Chem.* **275,** 30855–30863 (2000).

126. Jansen, G. *et al.* A structurally altered human reduced folate carrier with increased folic acid transport mediates a novel mechanism of antifolate resistance. *J. Biol. Chem.* **273,** 30189–98 (1998).

127. Gifford, A. J. *et al.* P-glycoprotein-mediated methotrexate resistance in CCRF-CEM sublines deficient in methotrexate accumulation due to a point mutation in the reduced folate carrier gene. *Int. J. cancer* **78,** 176–81 (1998).

128. Bram, E. *et al.* Mutant Gly482 and Thr482 ABCG2 mediate high-level resistance to lipophilic antifolates. *Cancer Chemother. Pharmacol.* **58,** 826–834 (2006).

129. Shafran, A. *et al.* ABCG2 Harboring the Gly482 Mutation Confers High-Level Resistance to Various Hydrophilic Antifolates. *Cancer Res.* **65,** 8414–8422 (2005).

130. Chen, Z.-S. *et al.* Analysis of methotrexate and folate transport by multidrug resistance protein 4 (ABCC4): MRP4 is a component of the methotrexate efflux system. *Cancer Res.* **62,** 3144–50 (2002).

131. Horns, R. C., Dower, W. J. & Schimke, R. T. Gene amplification in a leukemic patient treated with methotrexate. *J. Clin. Oncol.* **2,** 2–7 (1984).

132. Trent, J. M., Buick, R. N., Olson, S., Horns, R. C. & Schimke, R. T. Cytologic evidence for gene amplification in methotrexate-resistant cells obtained from a patient with ovarian adenocarcinoma. *J. Clin. Oncol.* **2,** 8–15 (1984).

133. Spencer, H. T. *et al.* Mutations in the gene for human dihydrofolate reductase: an unlikely cause of clinical relapse in pediatric leukemia after therapy with methotrexate. *Leukemia* **10,** 439–46 (1996).

134. Wang, L. G., Liu, X. M., Kreis, W. & Budman, D. R. The effect of antimicrotubule agents on signal transduction pathways of apoptosis: a review. *Cancer Chemother. Pharmacol.* **44,** 355–361 (1999).

135. Drake, K. J. *et al.* RALBP1/RLIP76 mediates multidrug resistance. *Int. J. Oncol.* **30,** 139–44 (2007).

136. Hopper-Borge, E. *et al.* Human Multidrug Resistance Protein 7 (ABCC10) Is a Resistance Factor for Nucleoside Analogues and Epothilone B. *Cancer Res.* **69,** 178–184 (2009).

137. Huang, R., Murry, D. J., Kolwankar, D., Hall, S. D. & Foster, D. R. Vincristine transcriptional regulation of efflux drug transporters in carcinoma cell lines. *Biochem. Pharmacol.* **71,** 1695–1704 (2006).

138. Koh, G. S., Jiang, N., Kham, S. K. Y., Chew, F. T. & Yeoh, A. E. J. Abstract 822: Drug resistance towards vincristine in acute lymphoblastic leukemia is mediated by the PI3K-Akt pathway. *Cancer Res.* **72,** 822–822 (2012).

139. Imai, Y. *et al.* Breast cancer resistance protein/ABCG2 is differentially regulated downstream of extracellular signal-regulated kinase. *Cancer Sci.* **100,** 1118–27 (2009).

140. Moore, M. J. Clinical Pharmacokinetics of Cyclophosphamide. *Clin. Pharmacokinet.* **20,** 194–208 (1991).

141. Crom, W. R. *et al.* Pharmacokinetics of Anticancer Drugs in Children. *Clin. Pharmacokinet.* **12,** 168–213 (1987).

142. Hall, A. G. & Tilby, M. J. *Mechanisms of Action of, and Modes of Resistance to, Alkylating Agents Used in the Treatment of Haematological Malignancies*.

143. Sarkaria, J. N. *et al.* Mechanisms of Chemoresistance to Alkylating Agents in Malignant Glioma. *Clin. Cancer Res.* **14,** 2900–2908 (2008).

144. Dimeloe, S. *et al.* Human regulatory T cells lack the cyclophosphamide-extruding transporter ABCB1 and are more susceptible to cyclophosphamide-induced apoptosis. *Eur. J. Immunol.* **44,** 3614–3620 (2014).

145. Beum, P. V, Kennedy, A. D., Williams, M. E., Lindorfer, M. A. & Taylor, R. P. The shaving reaction: rituximab/CD20 complexes are removed from mantle cell lymphoma and chronic lymphocytic leukemia cells by THP-1 monocytes. *J. Immunol.* **176,** 2600–9 (2006).

146. Mirski, S. E., Gerlach, J. H. & Cole, S. P. Multidrug resistance in a human small cell lung cancer cell line selected in adriamycin. *Cancer Res.* **47,** 2594–8 (1987).

147. Doyle, L. A. *et al.* A multidrug resistance transporter from human MCF-7 breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 15665–70 (1998).

148. Allikmets, R., Schriml, L. M., Hutchinson, A., Romano-Spica, V. & Dean, M. A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance. *Cancer Res.* **58,** 5337–9 (1998).

149. Miyake, K. *et al.* Molecular cloning of cDNAs which are highly overexpressed in mitoxantrone-resistant cells: demonstration of homology to ABC transport genes. *Cancer Res.* **59,** 8–13 (1999).

150. Leonard, G. D., Fojo, T. & Bates, S. E. The role of ABC transporters in clinical practice. *Oncologist* **8,** 411–24 (2003).

151. Gottesman, M. M., Fojo, T. & Bates, S. E. Multidrug resistance in cancer: role of ATP–dependent transporters. *Nat. Rev. Cancer* **2,** 48–58 (2002).

152. Leonard, G. D., Polgar, O. & Bates, S. E. ABC transporters and inhibitors: new targets, new agents. *Curr. Opin. Investig. Drugs* **3,** 1652–9 (2002).

153. Binkhathlan, Z. & Lavasanifar, A. P-glycoprotein inhibition as a therapeutic approach for overcoming multidrug resistance in cancer: current status and future perspectives. *Curr. Cancer Drug Targets* **13,** 326–46 (2013).

154. List, A. F. *et al.* Benefit of cyclosporine modulation of drug resistance in patients with poor-risk acute myeloid leukemia: a Southwest Oncology Group study. *Blood* **98,** 3212–20 (2001).

155. Yuan, Y. *et al.* Nanoparticle delivery of anticancer drugs overcomes multidrug resistance in breast cancer. *Drug Deliv.* **23,** 3350–3357 (2016).

156. Leget, G. A. & Czuczman, M. S. Use of rituximab, the new FDA-approved antibody. *Curr. Opin. Oncol.* **10,** 548–51 (1998).

157. Treon, S. P. *et al.* Tumor cell expression of CD59 is associated with resistance to CD20 serotherapy in patients with B-cell malignancies. *J. Immunother.* **24,** 263–71

158. Bannerji, R. *et al.* Apoptotic-Regulatory and Complement-Protecting Protein Expression in Chronic Lymphocytic Leukemia: Relationship to In Vivo Rituximab Resistance. *J. Clin. Oncol.* **21,** 1466–1471 (2003).

159. Manches, O. *et al.* In vitro mechanisms of action of rituximab on primary non-Hodgkin lymphomas. *Blood* **101,** 949–954 (2003).

160. Dall'Ozzo, S. *et al.* Rituximab-Dependent Cytotoxicity by Natural Killer Cells. *Cancer Res.* **64,** 4664–4669 (2004).

161. Lefebvre, M.-L., Krause, S. W., Salcedo, M. & Nardin, A. Ex Vivo-activated Human Macrophages Kill Chronic Lymphocytic Leukemia Cells in the Presence of Rituximab: Mechanism of Antibody-dependent Cellular Cytotoxicity and Impact of Human Serum. *J. Immunother.* **29,** 388–397 (2006).

162. Bowles, J. A. *et al.* Anti-CD20 monoclonal antibody with enhanced affinity for CD16 activates NK cells at lower concentrations and more effectively than rituximab. *Blood* **108,** 2648–2654 (2006).

163. Reff, M. E. *et al.* Depletion of B cells in vivo by a chimeric mouse human monoclonal antibody to CD20. *Blood* **83,** 435–45 (1994).

164. Flieger, D., Renoth, S., Beier, I., Sauerbruch, T. & Schmidt-Wolf, I. Mechanism of Cytotoxicity Induced by Chimeric Mouse Human Monoclonal Antibody IDEC-C2B8 in CD20-Expressing Lymphoma Cell Lines. *Cell. Immunol.* **204,** 55–63 (2000).

165. Hernandez-Ilizaliturri, F. J. *et al.* Neutrophils contribute to the biological antitumor activity of rituximab in a non-Hodgkin's lymphoma severe combined immunodeficiency mouse model. *Clin. Cancer Res.* **9,** 5866–73 (2003).

166. Stockmeyer, B. *et al.* Triggering Fc alpha-receptor I (CD89) recruits neutrophils as effector cells for CD20-directed antibody therapy. *J. Immunol.* **165,** 5954–61 (2000).

167. Koene, H. R. *et al.* Fc gammaRIIIa-158V/F polymorphism influences the binding of IgG by natural killer cell Fc gammaRIIIa, independently of the Fc gammaRIIIa-48L/R/H phenotype. *Blood* **90,** 1109–14 (1997).

168. Cartron, G. *et al.* Therapeutic activity of humanized anti-CD20 monoclonal antibody and polymorphism in IgG Fc receptor FcgammaRIIIa gene. *Blood* **99,** 754–8 (2002).

169. Weng, W.-K. & Levy, R. Two Immunoglobulin G Fragment C Receptor Polymorphisms Independently Predict Response to Rituximab in Patients With Follicular Lymphoma. *J. Clin. Oncol.* **21,** 3940–3947 (2003).

170. Bonavida, B. What signals are generated by anti-CD20 antibody therapy? *Curr. Hematol. Malig. Rep.* **1,** 205–213 (2006).

171. Jazirehi, A. R., Gan, X.-H., De Vos, S., Emmanouilides, C. & Bonavida, B. Rituximab (anti-CD20) selectively modifies Bcl-xL and apoptosis protease activating factor-1 (Apaf-1) expression and sensitizes human non-Hodgkin's lymphoma B cell lines to paclitaxel-induced apoptosis. *Mol. Cancer Ther.* **2,** 1183–93 (2003).

172. Ghosh, S. & Karin, M. Missing pieces in the NF-kappaB puzzle. *Cell* **109 Suppl,** S81-96 (2002).

173. Macor, P. *et al.* In vivo Targeting of Human Neutralizing Antibodies against CD55 and CD59 to Lymphoma Cells Increases the Antitumor Activity of Rituximab. *Cancer Res.* **67,** 10556–10563 (2007).

174. Kennedy, G. A. *et al.* Incidence and nature of CD20-negative relapses following rituximab therapy in aggressive B-cell non-Hodgkin's lymphoma: a retrospective review. *Br. J. Haematol.* **119,** 412–6 (2002).

175. Haidar, J. H. *et al.* Loss of CD20 expression in relapsed lymphomas after rituximab therapy. *Eur. J. Haematol.* **70,** 330–2 (2003).

176. Jazirehi, A. R., Vega, M. I. & Bonavida, B. Development of rituximab-resistant lymphoma clones with altered cell signaling and cross-resistance to chemotherapy. *Cancer Res.* **67,** 1270–81 (2007).

177. Czuczman, M. S. *et al.* Acquirement of rituximab resistance in lymphoma cell lines is associated with both global CD20 gene and protein down-regulation regulated at the pretranscriptional and posttranscriptional levels. *Clin. Cancer Res.* **14,** 1561–70 (2008).

178. Takei, K., Yamazaki, T., Sawada, U., Ishizuka, H. & Aizawa, S. Analysis of changes in CD20, CD55, and CD59 expression on established rituximab-resistant B-lymphoma cell lines. *Leuk. Res.* **30,** 625–631 (2006).

179. Vega, M. I. *et al.* Rituximab-Mediated Cell Signaling and Chemo/Immuno-sensitization of Drug-Resistant B-NHL Is Independent of Its Fc Functions. *Clin. Cancer Res.* **15,** 6582–6594 (2009).

180. Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15,** 81–94 (2017).

181. Anker, H. S. Biological Aspects of Cancer by Julian Huxley. *Perspect. Biol. Med.* **2,** 246–246 (1959).

182. NOWELL, P. C. & HUNGERFORD, D. A. Chromosome studies on normal and leukemic human leukocytes. *J. Natl. Cancer Inst.* **25,** 85–109 (1960).

183. Groffen, J. *et al.* Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22. *Cell* **36,** 93–9 (1984).

184. Biernaux, C., Sels, A., Huez, G. & Stryckmans, P. Very low level of major BCR-ABL expression in blood of some healthy individuals. *Bone Marrow Transplant.* **17 Suppl 3,** S45-7 (1996).

185. Biernaux, C., Loos, M., Sels, A., Huez, G. & Stryckmans, P. Detection of major bcr-abl gene expression at a very low level in blood cells of some healthy individuals. *Blood* **86,** 3118–22 (1995).

186. Bose, S., Deininger, M., Gora-Tybor, J., Goldman, J. M. & Melo, J. V. The presence of typical and atypical BCR-ABL fusion genes in leukocytes of normal individuals: biologic significance and implications for the assessment of minimal residual disease. *Blood* **92,** 3362–7 (1998).

187. Gerlinger, M. *et al.* Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *N. Engl. J. Med.* **366,** 883–892 (2012).

188. Fidler, I. J. & Kripke, M. L. Metastasis results from preexisting variant cells within a malignant tumor. *Science* **197,** 893–5 (1977).

189. Fidler, I. J. & Hart, I. R. Biological diversity in metastatic neoplasms: origins and implications. *Science* **217,** 998–1003 (1982).

190. Heppner, G. H. Tumor heterogeneity. *Cancer Res.* **44,** 2259–65 (1984).

191. Nowell, P. C. Mechanisms of tumor progression. *Cancer Res.* **46,** 2203–7 (1986).

192. Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194,** 23–8 (1976).

193. Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome — biological and translational implications. *Nat. Rev. Cancer* **11,** 726–734 (2011).

194. Nguyen, L. V., Vanner, R., Dirks, P. & Eaves, C. J. Cancer stem cells: an evolving concept. *Nat. Rev. Cancer* **12,** 133–143 (2012).

195. Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481,** 306–313 (2012).

196. Sakr, W. A., Haas, G. P., Cassin, B. F., Pontes, J. E. & Crissman, J. D. The frequency of carcinoma and intraepithelial neoplasia of the prostate in young male patients. *J. Urol.* **150,** 379–85 (1993).

197. Lapidot, T. *et al.* A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367,** 645–648 (1994).

198. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci.* **100,** 3983–3988 (2003).

199. Magee, J. A., Piskounova, E. & Morrison, S. J. Cancer Stem Cells: Impact, Heterogeneity, and Uncertainty. *Cancer Cell* **21,** 283–296 (2012).

200. Clarke, M. F. *et al.* Cancer Stem Cells—Perspectives on Current Status and Future Directions: AACR Workshop on Cancer Stem Cells. *Cancer Res.* **66,** 9339–9344 (2006).

201. Cohnheim, J. Congenitales, quergestreiftes Muskelsarkom der Nieren. *Virchows Arch.* **65,** 64–69 (1875).

202. Kreso, A. & Dick, J. E. Evolution of the Cancer Stem Cell Model. *Cell Stem Cell* **14,** 275–291 (2014).

203. Martinez-Climent, J. A., Fontan, L., Gascoyne, R. D., Siebert, R. & Prosper, F. Lymphoma stem cells: enough evidence to support their existence? *Haematologica* **95,** 293–302 (2010).

204. Bomken, S., Fišer, K., Heidenreich, O. & Vormoor, J. Understanding the cancer stem cell. *Br. J. Cancer* **103,** 439–445 (2010).

205. Gupta, P. B., Chaffer, C. L. & Weinberg, R. A. Cancer stem cells: mirage or reality? *Nat. Med.* **15,** 1010–1012 (2009).

206. Michor, F. & Polyak, K. The origins and implications of intratumor heterogeneity. *Cancer Prev. Res. (Phila).* **3,** 1361–4 (2010).

207. Cabrera, M. C., Hollingsworth, R. E. & Hurt, E. M. Cancer stem cell plasticity and tumor hierarchy. *World J. Stem Cells* **7,** 27 (2015).

208. Andriani, F. *et al.* Conversion to stem-cell state in response to microenvironmental cues is regulated by balance between epithelial and mesenchymal features in lung cancer cells. *Mol. Oncol.* **10,** 253–271 (2016).

209. Laplane, L. Cancer stem cells : philosophy and therapies. *Harvard University Press* (2016).

210. Chaffer, C. L. *et al.* Poised Chromatin at the ZEB1 Promoter Enables Breast Cancer Cell Plasticity and Enhances Tumorigenicity. *Cell* **154,** 61–74 (2013).

211. Chaffer, C. L. *et al.* Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 7950–5 (2011).

212. Quintana, E. *et al.* Phenotypic Heterogeneity among Tumorigenic Melanoma Cells from Patients that Is Reversible and Not Hierarchically Organized. *Cancer Cell* **18,** 510–523 (2010).

213. Roesch, A. *et al.* A Temporarily Distinct Subpopulation of Slow-Cycling Melanoma Cells Is Required for Continuous Tumor Growth. *Cell* **141,** 583–594 (2010).

214. Gupta, P. B. *et al.* Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells. *Cell* **146,** 633–644 (2011).

215. Suvà, M. L. *et al.* Reconstructing and Reprogramming the Tumor-Propagating Potential of Glioblastoma Stem-like Cells. *Cell* **157,** 580–594 (2014).

216. Sharma, S. V *et al.* A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* **141,** 69–80 (2010).

217. Shackleton, M., Quintana, E., Fearon, E. R. & Morrison, S. J. Heterogeneity in Cancer: Cancer Stem Cells versus Clonal Evolution. *Cell* **138,** 822–829 (2009).

218. Schwitalla, S. *et al.* Intestinal Tumorigenesis Initiated by Dedifferentiation and Acquisition of Stem-Cell-like Properties. *Cell* **152,** 25–38 (2013).

219. Suva, M. L., Riggi, N. & Bernstein, B. E. Epigenetic Reprogramming in Cancer. *Science (80-.).* **339,** 1567–1570 (2013).

220. Xie, H., Ye, M., Feng, R. & Graf, T. Stepwise reprogramming of B cells into macrophages. *Cell* **117,** 663–76 (2004).

221. Hanna, J. *et al.* Direct Reprogramming of Terminally Differentiated Mature B Lymphocytes to Pluripotency. *Cell* **133,** 250–264 (2008).

222. Gilbertson, R. J. & Rich, J. N. Making a tumour's bed: glioblastoma stem cells and the vascular niche. *Nat. Rev. Cancer* **7,** 733–736 (2007).

223. Charles, N. *et al.* Perivascular nitric oxide activates notch signaling and promotes stem-like character in PDGF-induced glioma cells. *Cell Stem Cell* **6,** 141–52 (2010).

224. Beck, B. *et al.* A vascular niche and a VEGF–Nrp1 loop regulate the initiation and stemness of skin tumours. *Nature* **478,** 399–403 (2011).

225. Kryczek, I. *et al.* IL-22+CD4+ T Cells Promote Colorectal Cancer Stemness via STAT3 Transcription Factor Activation and Induction of the Methyltransferase DOT1L. *Immunity* **40,** 772–784 (2014).

226. Chung, S. S. *et al.* Proinflammatory Cytokines IL-6 and TNF-α Increased Telomerase Activity through NF-κB/STAT1/STAT3 Activation, and Withaferin A Inhibited the Signaling in Colorectal Cancer Cells. *Mediators Inflamm.* **2017,** 5958429 (2017).

227. Shang, G.-S., Liu, L. & Qin, Y.-W. IL-6 and TNF-α promote metastasis of lung cancer

by inducing epithelial-mesenchymal transition. *Oncol. Lett.* **13,** 4657–4660 (2017).

228. Sansone, P. *et al.* IL-6 triggers malignant features in mammospheres from human ductal breast carcinoma and normal mammary gland. *J. Clin. Invest.* **117,** 3988–4002 (2007).

229. Zhang, M. *et al.* Intratumoral Heterogeneity in a Trp53-Null Mouse Model of Human Breast Cancer. *Cancer Discov.* **5,** 520–533 (2015).

230. Koch, R. *et al.* Populational equilibrium through exosome-mediated Wnt signaling in tumor progression of diffuse large B-cell lymphoma. *Blood* **123,** 2189–98 (2014).

231. Luckey, C. J. *et al.* Memory T and memory B cells share a transcriptional program of self-renewal with long-term hematopoietic stem cells. *Proc. Natl. Acad. Sci.* **103,** 3304–3309 (2006).

232. Lwin, T. *et al.* A microenvironment-mediated c-Myc/miR-548m/HDAC6 amplification loop in non-Hodgkin B cell lymphomas. *J. Clin. Invest.* **123,** 4612–4626 (2013).

233. Zhang, X. *et al.* Coordinated Silencing of MYC-Mediated miR-29 by HDAC3 and EZH2 as a Therapeutic Target of Histone Modification in Aggressive B-Cell Lymphomas. *Cancer Cell* **22,** 506–523 (2012).

234. Rappa, G. *et al.* Growth of cancer cell lines under stem cell-like conditions has the potential to unveil therapeutic targets. *Exp. Cell Res.* **314,** 2110–2122 (2008).

235. Wilson, A. *et al.* c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. *Genes Dev.* **18,** 2747–2763 (2004).

236. Koury, J., Zhong, L. & Hao, J. Targeting Signaling Pathways in Cancer Stem Cells for Cancer Treatment. *Stem Cells Int.* **2017,** 1–10 (2017).

237. Matsui, W. H. Cancer stem cell signaling pathways. *Medicine (Baltimore).* **95,** S8–S19 (2016).

238. Komiya, Y. & Habas, R. Wnt signal transduction pathways. *Organogenesis* **4,** 68–75 (2008).

239. Basu, S., Haase, G. & Ben-Ze'ev, A. Wnt signaling in cancer stem cells and colon cancer metastasis. *F1000Research* **5,** 699 (2016).

240. Mukherjee, S. *et al.* Curcumin inhibits breast cancer stem cell migration by amplifying the E-cadherin/β-catenin negative feedback loop. *Stem Cell Res. Ther.* **5,** 116 (2014).

241. Reya, T. *et al.* A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature* **423,** 409–414 (2003).

242. Zhang, X. *et al.* Wnt signaling regulates the stemness of lung cancer stem cells and its inhibitors exert anticancer effect on lung cancer SPC-A1 cells. *Med. Oncol.* **32,** 95 (2015).

243. Malanchi, I. *et al.* Cutaneous cancer stem cell maintenance is dependent on β-catenin signalling. *Nature* **452,** 650–653 (2008).

244. MacDonald, B. T., Tamai, K. & He, X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Dev. Cell* **17,** 9–26 (2009).

245. Medema, J. P. & Vermeulen, L. Microenvironmental regulation of stem cells in intestinal homeostasis and cancer. *Nature* **474,** 318–326 (2011).

246. Cubedo, E. *et al.* Identification of LMO2 transcriptome and interactome in diffuse large B-cell lymphoma. *Blood* **119,** 5478–5491 (2012).

247. Walther, N. *et al.* Aberrant Lymphocyte Enhancer–Binding Factor 1 Expression Is Characteristic for Sporadic Burkitt's Lymphoma. *Am. J. Pathol.* **182,** 1092–1098 (2013).

248. Clement, V., Sanchez, P., de Tribolet, N., Radovanovic, I. & Ruiz i Altaba, A. HEDGEHOG-GLI1 Signaling Regulates Human Glioma Growth, Cancer Stem Cell Self-Renewal, and Tumorigenicity. *Curr. Biol.* **17,** 165–172 (2007).

249. Bar, E. E. *et al.* Cyclopamine-Mediated Hedgehog Pathway Inhibition Depletes Stem-Like Cancer Cells in Glioblastoma. *Stem Cells* **25,** 2524–2533 (2007).

250. Dierks, C. *et al.* Expansion of Bcr-Abl-Positive Leukemic Stem Cells Is Dependent on Hedgehog Pathway Activation. *Cancer Cell* **14,** 238–249 (2008).

251. Zhao, C. *et al.* Hedgehog signalling is essential for maintenance of cancer stem cells in myeloid leukaemia. *Nature* **458,** 776–779 (2009).

252. Wang, F. *et al.* Hedgehog Signaling Regulates Epithelial-Mesenchymal Transition in Pancreatic Cancer Stem-Like Cells. *J. Cancer* **7,** 408–17 (2016).

253. Peacock, C. D. *et al.* Hedgehog signaling maintains a tumor stem cell compartment in multiple myeloma. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 4048–53 (2007).

254. Yoon, J. W. *et al.* Noncanonical Regulation of the Hedgehog Mediator *GLI1* by c-MYC in Burkitt Lymphoma. *Mol. Cancer Res.* **11,** 604–615 (2013).

255. Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch signaling: cell fate control and signal integration in development. *Science* **284,** 770–6 (1999).

256. Chen, Y. *et al.* Notch-1 signaling facilitates survivin expression in human non-small cell lung cancer cells. *Cancer Biol. Ther.* **11,** 14–21 (2011).

257. Yu, J.-B., Jiang, H. & Zhan, R.-Y. Aberrant Notch signaling in glioblastoma stem cells contributes to tumor recurrence and invasion. *Mol. Med. Rep.* **14,** 1263–1268 (2016).

258. Palomero, T. *et al.* NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. *Proc. Natl. Acad. Sci.* **103,** 18261–18266 (2006).

259. Zhong, Y. *et al.* NOTCH1 is a poor prognostic factor for breast cancer and is associated with breast cancer stem cells. *Onco. Targets. Ther.* **9,** 6865–6871 (2016).

260. Abel, E. V. *et al.* The Notch Pathway Is Important in Maintaining the Cancer Stem Cell Population in Pancreatic Cancer. *PLoS One* **9,** e91983 (2014).

261. Phillips, E. *et al.* Targeting atypical protein kinase C iota reduces viability in

glioblastoma stem-like cells *via* a notch signaling mechanism. *Int. J. Cancer* **139,** 1776–1787 (2016).

262. Cook, A. M. *et al.* Role of altered growth factor receptor-mediated JAK2 signaling in growth and maintenance of human acute myeloid leukemia stem cells. *Blood* **123,** 2826–2837 (2014).

263. Peñuelas, S. *et al.* TGF-β Increases Glioma-Initiating Cell Self-Renewal through the Induction of LIF in Human Glioblastoma. *Cancer Cell* **15,** 315–327 (2009).

264. Sherry, M. M., Reeves, A., Wu, J. K. & Cochran, B. H. STAT3 Is Required for Proliferation and Maintenance of Multipotency in Glioblastoma Stem Cells. *Stem Cells* **27,** 2383–2392 (2009).

265. Wang, X. *et al.* STAT3 mediates resistance of CD44+CD24−/low breast cancer stem cells to tamoxifen in vitro. *J. Biomed. Res.* **26,** 325–335 (2012).

266. CHUNG, S. S., GIEHL, N., WU, Y. & VADGAMA, J. V. STAT3 activation in HER2-overexpressing breast cancer promotes epithelial-mesenchymal transition and cancer stem cell traits. *Int. J. Oncol.* **44,** 403–411 (2014).

267. Tang, Y. *et al.* Progenitor/stem cells give rise to liver cancer due to aberrant TGF- and IL-6 signaling. *Proc. Natl. Acad. Sci.* **105,** 2445–2450 (2008).

268. Korkaya, H. *et al.* Activation of an IL6 Inflammatory Loop Mediates Trastuzumab Resistance in HER2+ Breast Cancer by Expanding the Cancer Stem Cell Population. *Mol. Cell* **47,** 570–584 (2012).

269. Johnstone, C. N., Chand, A., Putoczki, T. L. & Ernst, M. Emerging roles for IL-11 signaling in cancer development and progression: Focus on breast cancer. *Cytokine Growth Factor Rev.* **26,** 489–498 (2015).

270. Duan, S. *et al.* PTEN deficiency reprogrammes human neural stem cells towards a glioblastoma stem cell-like phenotype. *Nat. Commun.* **6,** 10068 (2015).

271. Lechman, E. R. *et al.* miR-126 Regulates Distinct Self-Renewal Outcomes in Normal and Malignant Hematopoietic Stem Cells. *Cancer Cell* **29,** 214–228 (2016).

272. Zhou, J. *et al.* Activation of the PTEN/mTOR/STAT3 pathway in breast cancer stem-like cells is required for viability and maintenance. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 16158–63 (2007).

273. Douville, J., Beaulieu, R. & Balicki, D. ALDH1 as a Functional Marker of Cancer Stem and Progenitor Cells. *Stem Cells Dev.* **18,** 17–26 (2009).

274. Huang, E. H. *et al.* Aldehyde Dehydrogenase 1 Is a Marker for Normal and Malignant Human Colonic Stem Cells (SC) and Tracks SC Overpopulation during Colon Tumorigenesis. *Cancer Res.* **69,** 3382–3389 (2009).

275. Huang, F.-F. *et al.* Inactivation of PTEN increases ABCG2 expression and the side population through the PI3K/Akt pathway in adult acute leukemia. *Cancer Lett.* **336,** 96–105 (2013).

276. Bleau, A.-M. *et al.* PTEN/PI3K/Akt pathway regulates the side population phenotype and ABCG2 activity in glioma tumor stem-like cells. *Cell Stem Cell* **4,** 226–35 (2009).

277. Jung, M.-J. *et al.* Upregulation of CXCR4 is functionally crucial for maintenance of stemness in drug-resistant non-small cell lung cancer cells. *Oncogene* **32,** 209–221 (2013).

278. Chang, W.-W. *et al.* The expression and significance of insulin-like growth factor-1 receptor and its pathway on breast cancer stem/progenitors. *Breast Cancer Res.* **15,** R39 (2013).

279. Liu, M. *et al.* The canonical NF-kappaB pathway governs mammary tumorigenesis in transgenic mice and tumor stem cell expansion. *Cancer Res.* **70,** 10464–73 (2010).

280. Chen, C. *et al.* IKKβ Enforces a LIN28B/TCF7L2 Positive Feedback Loop That Promotes Cancer Cell Stemness and Metastasis. *Cancer Res.* **75,** 1725–35 (2015).

281. Widera, D., Mikenberg, I., Elvers, M., Kaltschmidt, C. & Kaltschmidt, B. Tumor necrosis factor α triggers proliferation of adult neural stem cells via IKK/NF-κB signaling. *BMC Neurosci.* **7,** 64 (2006).

282. Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* **3,** 730–7 (1997).

283. Christian, B. *et al.* Mantle Cell Lymphoma 12 Years After Allogeneic Bone Marrow Transplantation Occurring Simultaneously in Recipient and Donor. *J. Clin. Oncol.* **28,** e629–e632 (2010).

284. Sala-Torra, O. *et al.* Evidence of Donor-Derived Hematologic Malignancies after Hematopoietic Stem Cell Transplantation. *Biol. Blood Marrow Transplant.* **12,** 511–517 (2006).

285. Hart, J. *et al.* Transmission of a Follicular Lymphoma by Allogeneic Bone Marrow Tranplantation — Evidence to Support the Existence of a Lymphoma Progenitor Cell. *Blood* **108,** (2006).

286. Weigert, O. *et al.* Molecular ontogeny of donor-derived follicular lymphomas occurring after hematopoietic cell transplantation. *Cancer Discov.* **2,** 47–55 (2012).

287. Janikova, A. *et al.* Transmission of t(11;14)-positive cells by allogeneic stem cell transplant: 10-year journey to mantle cell lymphoma. *Leuk. Lymphoma* **55,** 1935–1938 (2014).

288. Berg, K. D. *et al.* Transmission of a T-Cell Lymphoma by Allogeneic Bone Marrow Transplantation. *N. Engl. J. Med.* **345,** 1458–1463 (2001).

289. Jaiswal, S. *et al.* Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N. Engl. J. Med.* **371,** 2488–2498 (2014).

290. Genovese, G. *et al.* Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N. Engl. J. Med.* **371,** 2477–2487 (2014).

291. Asmar, F. *et al.* Genome-wide profiling identifies a DNA methylation signature that associates with TET2 mutations in diffuse large B-cell lymphoma. *Haematologica* **98,** 1912–1920 (2013).

292. Couronné, L., Bastard, C. & Bernard, O. A. *TET2* and *DNMT3A* Mutations in Human T-Cell Lymphoma. *N. Engl. J. Med.* **366,** 95–96 (2012).

293. Quivoron, C. *et al.* TET2 Inactivation Results in Pleiotropic Hematopoietic Abnormalities in Mouse and Is a Recurrent Event during Human Lymphomagenesis. *Cancer Cell* **20,** 25–38 (2011).

294. Quijada-Álamo, M. *et al.* Next-generation sequencing and FISH studies reveal the appearance of gene mutations and chromosomal abnormalities in hematopoietic progenitors in chronic lymphocytic leukemia. *J. Hematol. Oncol.* **10,** 83 (2017).

295. Chung, S. S. *et al.* Hematopoietic Stem Cell Origin of BRAFV600E Mutations in Hairy Cell Leukemia. *Sci. Transl. Med.* **6,** 238ra71-238ra71 (2014).

296. Sebastián, E. *et al.* High-resolution copy number analysis of paired normal-tumor samples from diffuse large B cell lymphoma. *Ann. Hematol.* **95,** 253–262 (2016).

297. Jan, M. *et al.* Absence of Evidence Implicating Hematopoietic Stem Cells As Common Progenitors for DLBCL Mutations. *Blood* **128,** (2016).

298. Martin-Subero, J. I. *et al.* New insights into the biology and origin of mature aggressive B-cell lymphomas by combined epigenomic, genomic, and transcriptional profiling. *Blood* **113,** 2488–2497 (2009).

299. Bossen, C. *et al.* Interactions of tumor necrosis factor (TNF) and TNF receptor family members in the mouse and human. *J. Biol. Chem.* **281,** 13964–71 (2006).

300. Rongvaux, A. *et al.* Human Hemato-Lymphoid System Mice: Current Use and Future Potential for Medicine. *Annu. Rev. Immunol.* **31,** 635–674 (2013).

301. Theocharides, A. P. A., Rongvaux, A., Fritsch, K., Flavell, R. A. & Manz, M. G. Humanized hemato-lymphoid system mice. *Haematologica* **101,** 5–19 (2016).

302. Shultz, L. D. *et al.* Multiple defects in innate and adaptive immunologic function in NOD/LtSz-scid mice. *J. Immunol.* **154,** 180–91 (1995).

303. Shultz, L. D. *et al.* Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *J. Immunol.* **174,** 6477–89 (2005).

304. Quintana, E. *et al.* Efficient tumour formation by single human melanoma cells. *Nature* **456,** 593–598 (2008).

305. Ishikawa, F. *et al.* Development of functional human blood and immune systems in NOD/SCID/IL2 receptor chainnull mice. *Blood* **106,** 1565–1573 (2005).

306. Mullen, P. The Use of Matrigel to Facilitate the Establishment of Human Cancer Cell Lines as Xenografts. in *Cancer Cell Culture* 287–292 (Humana Press, 2004). doi:10.1385/1-59259-406-9:287

307. Kleinman, H. K. *et al.* Basement membrane complexes with biological activity. *Biochemistry* **25,** 312–318 (1986).

308. Schober, M. & Fuchs, E. Tumor-initiating stem cells of squamous cell carcinomas and their control by TGF-β and integrin/focal adhesion kinase (FAK) signaling. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 10544–9 (2011).

309. Yeung, T. M., Gandhi, S. C., Wilding, J. L., Muschel, R. & Bodmer, W. F. Cancer stem cells from colorectal cancer-derived cell lines. *Proc. Natl. Acad. Sci.* **107,** 3722–3727 (2010).

310. Ricci-Vitiani, L. *et al.* Identification and expansion of human colon-cancer-initiating cells. *Nature* **445,** 111–115 (2007).

311. Eramo, A. *et al.* Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ.* **15,** 504–514 (2008).

312. Boiko, A. D. *et al.* Human melanoma-initiating cells express neural crest nerve growth factor receptor CD271. *Nature* **466,** 133–137 (2010).

313. Inoue, A. *et al.* Cancer stem-like cells of glioblastoma characteristically express MMP-13 and display highly invasive activity. *Int. J. Oncol.* **37,** 1121–31 (2010).

314. Qin, X. *et al.* Intrinsic gemcitabine resistance in a novel pancreatic cancer cell line is associated with cancer stem cell-like phenotype. *Int. J. Oncol.* **40,** 798–806 (2011).

315. Ginestier, C. *et al.* ALDH1 Is a Marker of Normal and Malignant Human Mammary Stem Cells and a Predictor of Poor Clinical Outcome. *Cell Stem Cell* **1,** 555–567 (2007).

316. Kim, M. P. *et al.* ALDH Activity Selectively Defines an Enhanced Tumor-Initiating Cell Population Relative to CD133 Expression in Human Pancreatic Adenocarcinoma. *PLoS One* **6,** e20636 (2011).

317. Baum, C. M., Weissman, I. L., Tsukamoto, A. S., Buckle, A. M. & Peault, B. Isolation of a candidate human hematopoietic stem-cell population. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 2804–8 (1992).

318. Taussig, D. C. *et al.* Anti-CD38 antibody-mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. *Blood* **112,** 568–75 (2008).

319. Eppert, K. *et al.* Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat. Med.* **17,** 1086–1093 (2011).

320. Cariati, M. *et al.* Alpha-6 integrin is necessary for the tumourigenicity of a stem cell-like subpopulation within the MCF7 breast cancer cell line. *Int. J. Cancer* **122,** 298–304 (2008).

321. Wright, M. H. *et al.* Brca1 breast tumors contain distinct CD44+/CD24- and CD133+cells with cancer stem cell characteristics. *Breast Cancer Res.* **10,** R10 (2008).

322. Chen, W. *et al.* Identification of CD44+ cancer stem cells in human gastric cancer. *Hepatogastroenterology.* **60,** 949–54 (2013).

323. Gao, M.-Q., Choi, Y.-P., Kang, S., Youn, J. H. & Cho, N.-H. CD24+ cells from hierarchically organized ovarian cancer are enriched in cancer stem cells. *Oncogene* **29,** 2672–2680 (2010).

324. Chau, W. K., Ip, C. K., Mak, A. S. C., Lai, H.-C. & Wong, A. S. T. c-Kit mediates chemoresistance and tumor-initiating capacity of ovarian cancer cells through activation of Wnt/β-catenin–ATP-binding cassette G2 signaling. *Oncogene* **32,** 2767–2781 (2013).

325. Curley, M. D. *et al.* CD133 Expression Defines a Tumor Initiating Cell Population in Primary Human Ovarian Cancer. *Stem Cells* **27,** N/A-N/A (2009).

326. Stewart, J. M. *et al.* Phenotypic heterogeneity and instability of human ovarian tumor-initiating cells. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 6468–73 (2011).

327. Gentner, B. *et al.* Identification of Hematopoietic Stem Cell-Specific miRNAs Enables Gene Therapy of Globoid Cell Leukodystrophy. *Sci. Transl. Med.* **2,** 58ra84-58ra84 (2010).

328. Vermeulen, L. *et al.* Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nat. Cell Biol.* **12,** 468–476 (2010).

329. Venter, J. C. *et al.* The Sequence of the Human Genome. *Science (80-.).* **291,** 1304–1351 (2001).

330. Robey, R. W. *et al.* ABCG2: a perspective. *Adv. Drug Deliv. Rev.* **61,** 3–13 (2009).

331. Schinkel. P-Glycoprotein, a gatekeeper in the blood-brain barrier. *Adv. Drug Deliv. Rev.* **36,** 179–194 (1999).

332. Sun, H., Dai, H., Shaik, N. & Elmquist, W. F. Drug efflux transporters in the CNS. *Adv. Drug Deliv. Rev.* **55,** 83–105 (2003).

333. Zhou, S. *et al.* The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. *Nat. Med.* **7,** 1028–1034 (2001).

334. Goodell, M. A., Brose, K., Paradis, G., Conner, A. S. & Mulligan, R. C. Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. *J. Exp. Med.* **183,** 1797–806 (1996).

335. Camargo, F. D., Chambers, S. M., Drew, E., McNagny, K. M. & Goodell, M. A. Hematopoietic stem cells do not engraft with absolute efficiencies. *Blood* **107,** 501–507 (2006).

336. Challen, G. A. & Little, M. H. A Side Order of Stem Cells: The SP Phenotype. *Stem Cells* **24,** 3–12 (2006).

337. Storms, R. W., Goodell, M. A., Fisher, A., Mulligan, R. C. & Smith, C. Hoechst dye efflux reveals a novel CD7+CD34− lymphoid progenitor in human umbilical cord blood. *Blood* **96,** (2000).

338. Asakura, A., Seale, P., Girgis-Gabardo, A. & Rudnicki, M. A. Myogenic specification of side population cells in skeletal muscle. *J. Cell Biol.* **159,** 123–134 (2002).

339. Iwatani, H. *et al.* Hematopoietic and nonhematopoietic potentials of Hoechstlow/side population cells isolated from adult rat kidney. *Kidney Int.* **65,** 1604–1614 (2004).

340. Shimano, K. *et al.* Hepatic oval cells have the side population phenotype defined by expression of ATP-binding cassette transporter ABCG2/BCRP1. *Am. J. Pathol.* **163,** 3–9 (2003).

341. Clayton, H., Titley, I. & Vivanco, M. dM. Growth and differentiation of progenitor/stem cells derived from the human mammary gland. *Exp. Cell Res.* **297,** 444–460 (2004).

342. Summer, R. *et al.* Side population cells and Bcrp1 expression in lung. *Am. J. Physiol. Cell. Mol. Physiol.* **285,** L97–L104 (2003).

343. Chua, C. *et al.* Characterization of a side population of astrocytoma cells in response to temozolomide. *J. Neurosurg.* **109,** 856–866 (2008).

344. Dean, M., Fojo, T. & Bates, S. Tumour stem cells and drug resistance. *Nat. Rev. Cancer* **5,** 275–284 (2005).

345. Hirschmann-Jax, C. *et al.* A distinct "side population" of cells with high drug efflux capacity in human tumor cells. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 14228–33 (2004).

346. Ling, V. P-glycoprotein: its role in drug resistance. *Am. J. Med.* **99,** 31S–34S (1995).

347. Szotek, P. P. *et al.* Ovarian cancer side population defines cells with stem cell-like characteristics and Mullerian Inhibiting Substance responsiveness. *Proc. Natl. Acad. Sci.* **103,** 11154–11159 (2006).

348. Wu, C. *et al.* Side Population Cells Isolated from Mesenchymal Neoplasms Have Tumor Initiating Potential. *Cancer Res.* **67,** 8216–8222 (2007).

349. Li, R., Wu, X., Wei, H. & Tian, S. Characterization of side population cells isolated from the gastric cancer cell line SGC-7901. *Oncol. Lett.* **5,** 877–883 (2013).

350. Triel, C., Vestergaard, M. E., Bolund, L., Jensen, T. G. & Jensen, U. B. Side population cells in human and mouse epidermis lack stem cell characteristics. *Exp. Cell Res.* **295,** 79–90 (2004).

351. Zeng, H. *et al.* Lack of ABCG2 Expression and Side Population Properties in Human Pluripotent Stem Cells. *Stem Cells* **27,** 2435–2445 (2009).

352. Mayer, U. *et al.* Substantial excretion of digoxin via the intestinal mucosa and prevention of long-term digoxin accumulation in the brain by the mdr 1a P-glycoprotein. *Br. J. Pharmacol.* **119,** 1038–44 (1996).

353. Smit, J. W., Schinkel, A. H., Weert, B. & Meijer, D. K. Hepatobiliary and intestinal clearance of amphiphilic cationic drugs in mice in which both mdr1a and mdr1b genes have been disrupted. *Br. J. Pharmacol.* **124,** 416–24 (1998).

354. Choo, E. F. *et al.* Pharmacological inhibition of P-glycoprotein transport enhances the distribution of HIV-1 protease inhibitors into brain and testes. *Drug Metab. Dispos.* **28,** 655–60 (2000).

355. Lankas, G. R., Wise, L. D., Cartwright, M. E., Pippert, T. & Umbenhauer, D. R. Placental P-glycoprotein deficiency enhances susceptibility to chemically induced birth defects in mice. *Reprod. Toxicol.* **12,** 457–63

356. Golebiewska, A., Brons, N. H. C., Bjerkvig, R. & Niclou, S. P. Critical Appraisal of the Side Population Assay in Stem Cell and Cancer Stem Cell Research. *Cell Stem Cell* **8,** 136–147 (2011).

357. Natarajan, K., Xie, Y., Baer, M. R. & Ross, D. D. Role of breast cancer resistance protein (BCRP/ABCG2) in cancer drug resistance. *Biochem. Pharmacol.* **83,** 1084–1103 (2012).

358. Malfuson, J.-V. *et al.* SP/drug efflux functionality of hematopoietic progenitors is controlled by mesenchymal niche through VLA-4/CD44 axis. *Leukemia* **28,** 853–864 (2014).

359. Kim, J. Bin *et al.* Side population in LX2 cells decreased by transforming growth factor-β. *Hepatol. Res.* **44,** 229–237 (2014).

360. Liu, P.-P. *et al.* Metabolic regulation of cancer cell side population by glucose through activation of the Akt pathway. *Cell Death Differ.* **21,** 124–35 (2014).

361. To, K. K. W. *et al.* Escape from hsa-miR-519c enables drug-resistant cells to maintain high expression of ABCG2. *Mol. Cancer Ther.* **8,** 2959–68 (2009).

362. Wang, F. *et al.* hsa-miR-520h downregulates ABCG2 in pancreatic cancer cells to inhibit migration, invasion, and side populations. *Br. J. Cancer* **103,** 567–74 (2010).

363. Pan, Y.-Z., Morris, M. E. & Yu, A.-M. MicroRNA-328 Negatively Regulates the Expression of Breast Cancer Resistance Protein (BCRP/ABCG2) in Human Cancer Cells. *Mol. Pharmacol.* **75,** 1374–1379 (2009).

364. Motulsky, H. & Christopoulos, A. Fitting Models to Biological Data using Linear and Nonlinear Regression Contents at a Glance. *Oxford Univ. Press* (2014).

365. Foucquier, J. & Guedj, M. Analysis of drug combinations: current methodological landscape. *Pharmacol. Res. Perspect.* **3,** e00149 (2015).

366. Wang, X., Spandidos, A., Wang, H. & Seed, B. PrimerBank: a PCR primer database for quantitative gene expression analysis, 2012 update. *Nucleic Acids Res.* **40,** D1144–D1149 (2012).

367. PrimerBank.

368. Simon Andrews. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

369. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* **46,** W537–W544 (2018).

370. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12,** 357–60 (2015).

371. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).

372. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 550 (2014).

373. Metsalu, T. & Vilo, J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **43,** W566–W570 (2015).

374. Ni, Y. *et al.* Gene set enrichment analysis: A genome-wide expression profile-based strategy for discovering functional microRNA–disease relationships. *J. Int. Med. Res.* **46,** 596–611 (2018).

375. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517,** 583–588 (2014).

376. Joung, J. *et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12,** 828–863 (2017).

377. Nishimasu, H. *et al.* Crystal Structure of Cas9 in Complex with Guide RNA and

Target DNA. *Cell* **156,** 935–949 (2014).

378. Cas9 Activator Tool. http://sam.genome-engineering.org/database/

379. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17,** 10–12 (2011).

380. Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15,** 554 (2014).

381. Li, W. *et al.* Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol.* **16,** 281 (2015).

382. Daniel, V. C. *et al.* A Primary Xenograft Model of Small-Cell Lung Cancer Reveals Irreversible Changes in Gene Expression Imposed by Culture In vitro. *Cancer Res.* **69,** 3364–3373 (2009).

383. Sausville, E. A. & Burger, A. M. Contributions of Human Tumor Xenografts to Anticancer Drug Development. *Cancer Res.* **66,** 3351–3354 (2006).

384. Uzunova, L. & Burke, A. Update on non-Hodgkin lymphoma in children. *Paediatr. Child Health (Oxford).* **26,** 57–62 (2016).

385. Chapuy, B. *et al.* Diffuse large B-cell lymphoma patient-derived xenograft models capture the molecular and biological heterogeneity of the disease. *Blood* **127,** 2203–2213 (2016).

386. DeRose, Y. S. *et al.* Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nat. Med.* **17,** 1514–1520 (2011).

387. Zhao, X. *et al.* Global gene expression profiling confirms the molecular fidelity of primary tumor-based orthotopic xenograft mouse models of medulloblastoma. *Neuro. Oncol.* **14,** 574–583 (2012).

388. McEvoy, J. *et al.* Analysis of MDM2 and MDM4 Single Nucleotide Polymorphisms, mRNA Splicing and Protein Expression in Retinoblastoma. *PLoS One* **7,** e42739 (2012).

389. Reyal, F. *et al.* Molecular profiling of patient-derived breast cancer xenografts. *Breast Cancer Res.* **14,** R11 (2012).

390. Ding, L. *et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464,** 999–1005 (2010).

391. Bergamaschi, A. *et al.* Molecular profiling and characterization of luminal-like and basal-like *in vivo* breast cancer xenograft models. *Mol. Oncol.* **3,** 469–482 (2009).

392. Zhang, L. *et al.* Personalized Medicine and Imaging B-Cell Lymphoma Patient-Derived Xenograft Models Enable Drug Discovery and Are a Platform for Personalized Therapy. (2017). doi:10.1158/1078-0432.CCR-16-2703

393. Voskoglou-Nomikos, T. *et al.* Clinical predictive value of the in vitro cell line, human xenograft, and mouse allograft preclinical cancer models. *Clin. Cancer Res.* **9,** 4227–39 (2003).

394. Naeim, F. *et al.* Principles of Immunophenotyping. *Atlas Hematop.* 25–46 (2013). doi:10.1016/B978-0-12-385183-3.00002-4

395. Dobbin, Z. C. *et al.* Using heterogeneity of the patient-derived xenograft model to identify the chemoresistant population in ovarian cancer. *Oncotarget* **5,** 8750–8764 (2014).

396. Cozma, D. *et al.* B cell activator PAX5 promotes lymphomagenesis through stimulation of B cell receptor signaling. *J. Clin. Invest.* **117,** 2602–10 (2007).

397. Berney, T. *et al.* Patterns of engraftment in different strains of immunodeficient mice reconstituted with human peripheral blood lymphocytes. *Transplantation* **72,** 133–40 (2001).

398. Agliano, A. *et al.* Human acute leukemia cells injected in NOD/LtSz- *scid/IL-2Rγ* null mice generate a faster and more efficient disease compared to other NOD/ *scid* -related strains. *Int. J. Cancer* **123,** 2222–2227 (2008).

399. Siolas, D. & Hannon, G. J. Patient-derived tumor xenografts: transforming clinical samples into mouse models. *Cancer Res.* **73,** 5315–9 (2013).

400. Scott, D. W. & Gascoyne, R. D. The tumour microenvironment in B cell lymphomas. *Nat. Rev. Cancer* **14,** 517–534 (2014).

401. Coupland, S. E. & Coupland, S. E. The challenge of the microenvironment in B-cell lymphomas. *Histopathology* **58,** 69–80 (2011).

402. Shain, K. H., Dalton, W. S. & Tao, J. The tumor microenvironment shapes hallmarks of mature B-cell malignancies. *Oncogene* **34,** 4673–4682 (2015).

403. Shain, K. H. & Tao, J. The B-cell receptor orchestrates environment-mediated lymphoma survival and drug resistance in B-cell malignancies. *Oncogene* **33,** 4107–13 (2014).

404. Lwin, T. *et al.* Bone marrow stromal cells prevent apoptosis of lymphoma cells by upregulation of anti-apoptotic proteins associated with activation of NF-κB (RelB/p52) in non-Hodgkin's lymphoma cells. *Leukemia* **21,** 1521–1531 (2007).

405. Mitsiades, C. S. *et al.* The Role of the Bone Marrow Microenvironment in the Pathophysiology of Myeloma and Its Significance in the Development of More Effective Therapies. *Hematol. Oncol. Clin. North Am.* **21,** 1007–1034 (2007).

406. Medina, D. J. *et al.* Mesenchymal stromal cells protect mantle cell lymphoma cells from spontaneous and drug-induced apoptosis through secretion of B-cell activating factor and activation of the canonical and non-canonical nuclear factor κB pathways. *Haematologica* **97,** 1255–63 (2012).

407. Lichtenstein, A., Tu, Y., Fady, C., Vescio, R. & Berenson, J. Interleukin-6 Inhibits Apoptosis of Malignant Plasma Cells. *Cell. Immunol.* **162,** 248–255 (1995).

408. Gilbert, L. A. & Hemann, M. T. Context-specific roles for paracrine IL-6 in lymphomagenesis. *Genes Dev.* **26,** 1758–68 (2012).

409. Li, Y. *et al.* DF3/MUC1 signaling in multiple myeloma cells is regulated by interleukin-7. *Cancer Biol. Ther.* **2,** 187–93

410. Kimlinger, T. *et al.* Differential expression of vascular endothelial growth factors and their receptors in multiple myeloma. *Haematologica* **91,** 1033–40 (2006).

411. Möller, C., Strömberg, T., Juremalm, M., Nilsson, K. & Nilsson, G. Expression and

function of chemokine receptors in human multiple myeloma. *Leukemia* **17,** 203–210 (2003).

412. Jundt, F. *et al.* Jagged1-induced Notch signaling drives proliferation of multiple myeloma cells. *Blood* **103,** 3511–3515 (2004).

413. Hodge, L. S. *et al.* IL-21 in the bone marrow microenvironment contributes to IgM secretion and proliferation of malignant cells in Waldenstrom macroglobulinemia. *Blood* **120,** 3774–3782 (2012).

414. Hidalgo, M. *et al.* A pilot clinical study of treatment guided by personalized tumorgrafts in patients with advanced cancer. *Mol. Cancer Ther.* **10,** 1311–6 (2011).

415. Visvader, J. E. & Lindeman, G. J. Cell Stem Cell Perspective Cancer Stem Cells: Current Status and Evolving Complexities. (2012). doi:10.1016/j.stem.2012.05.007

416. Tysnes, B. B. Tumor-initiating and -propagating cells: cells that we would like to identify and control. *Neoplasia* **12,** 506–15 (2010).

417. Wicha, M. S., Liu, S. & Dontu, G. Cancer Stem Cells: An Old Idea—A Paradigm Shift. *Cancer Res.* **66,** 1883–1890 (2006).

418. Dieter, S. M. *et al.* Distinct Types of Tumor-Initiating Cells Form Human Colon Cancer Tumors and Metastases. *Cell Stem Cell* **9,** 357–365 (2011).

419. Zimmerlin, L., Donnenberg, V. S. & Donnenberg, A. D. Rare Event Detection and Analysis in Flow Cytometry: Bone Marrow Mesenchymal Stem Cells, Breast Cancer Stem/Progenitor Cells in Malignant Effusions, and Pericytes in Disaggregated Adipose Tissue. in *Methods in molecular biology (Clifton, N.J.)* **699,** 251–273 (2011).

420. Islam, F., Gopalan, V., Smith, R. A. & Lam, A. K.-Y. Translational potential of cancer stem cells: A review of the detection of cancer stem cells and their roles in cancer recurrence and cancer treatment. *Exp. Cell Res.* **335,** 135–147 (2015).

421. Zhang, C., Li, C., He, F., Cai, Y. & Yang, H. Identification of CD44+CD24+ gastric cancer stem cells. *J. Cancer Res. Clin. Oncol.* **137,** 1679–1686 (2011).

422. Yamazaki, H. *et al.* Regulation of cancer stem cell properties by CD9 in human B-acute lymphoblastic leukemia. *Biochem. Biophys. Res. Commun.* **409,** 14–21 (2011).

423. Lau, W. M. *et al.* CD44v8-10 Is a Cancer-Specific Marker for Gastric Cancer Stem Cells. *Cancer Res.* **74,** 2630–2641 (2014).

424. Thapa, R. & Wilson, G. D. The Importance of CD44 as a Stem Cell Biomarker and Therapeutic Target in Cancer. *Stem Cells Int.* **2016,** 1–15 (2016).

425. Blackburn, J. S., Liu, S. & Langenau, D. M. Quantifying the frequency of tumor-propagating cells using limiting dilution cell transplantation in syngeneic zebrafish. *J. Vis. Exp.* e2790 (2011). doi:10.3791/2790

426. Bierie, B. *et al.* Integrin-β4 identifies cancer stem cell-enriched populations of partially mesenchymal carcinoma cells. *Proc. Natl. Acad. Sci. U. S. A.* **114,** E2337–E2346 (2017).

427. Hu, Y. & Smyth, G. K. ELDA: Extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **347,** 70–78 (2009).

428. Doulatov, S., Notta, F., Laurenti, E. & Dick, J. E. Hematopoiesis: A Human Perspective. *Cell Stem Cell* **10,** 120–136 (2012).

429. Notta, F. *et al.* Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science* **333,** 218–21 (2011).

430. Moti, N. *et al.* Anaplastic large cell lymphoma-propagating cells are detectable by side population analysis and possess an expression profile reflective of a primitive origin. *Oncogene* **34,** 1843–1852 (2015).

431. GUO, Z. *et al.* Side population in hepatocellular carcinoma HCCLM3 cells is enriched with stem-like cancer cells. *Oncol. Lett.* **11,** 3145–3151 (2016).

432. Teshima, K. *et al.* Dysregulation of BMI1 and microRNA-16 collaborate to enhance an anti-apoptotic potential in the side population of refractory mantle cell lymphoma. *Oncogene* **33,** 2191–2203 (2014).

433. Ovalle, S. *et al.* The tetraspanin CD9 inhibits the proliferation and tumorigenicity of human colon carcinoma cells. *Int. J. Cancer* **121,** 2140–2152 (2007).

434. TANG, M. *et al.* Downregulation of CD9 promotes pancreatic cancer growth and metastasis through upregulation of epidermal growth factor on the cell surface. *Oncol. Rep.* **34,** 350–358 (2015).

435. Nishida, H. *et al.* CD9 correlates with cancer stem cell potentials in human B-acute lymphoblastic leukemia cells. *Biochem. Biophys. Res. Commun.* **382,** 57–62 (2009).

436. Yang, C.-H. *et al.* Identification of CD24 as a Cancer Stem Cell Marker in Human Nasopharyngeal Carcinoma. *PLoS One* **9,** e99412 (2014).

437. Sheridan, C. *et al.* CD44+/CD24-breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. *Breast Cancer Res.* **8,** R59 (2006).

438. Hirata, K. *et al.* CD44 variant 9 expression in primary early gastric cancer as a predictive marker for recurrence. *Br. J. Cancer* **109,** 379–86 (2013).

439. Ozawa, M. *et al.* Prognostic significance of CD44 variant 2 upregulation in colorectal cancer. *Br. J. Cancer* **111,** 365–74 (2014).

440. Roudi, R. *et al.* Clinical significance of putative cancer stem cell marker CD44 in different histological subtypes of lung cancer. *Cancer Biomark.* **14,** 457–67 (2014).

441. Miyatake, Y., Sheehy, N., Ikeshita, S., Hall, W. W. & Kasahara, M. Anchorage-dependent multicellular aggregate formation induces CD44 high cancer stem cell-like ATL cells in an NF-κB- and vimentin-dependent manner. *Cancer Lett.* **357,** 355–363 (2015).

442. Shanafelt, T. D. *et al.* CD49d expression is an independent predictor of overall survival in patients with chronic lymphocytic leukaemia: a prognostic parameter with therapeutic potential. *Br. J. Haematol.* **140,** 537–46 (2008).

443. Noh, D.-Y. *et al.* The CD49d+/high subpopulation from isolated human breast sarcoma spheres possesses tumor-initiating ability. *Int. J. Oncol.* **40,** 665–672 (2011).

444. Agematsu, K., Hokibara, S., Nagumo, H. & Komiyama, A. CD27: a memory B-cell marker. *Immunol. Today* **21,** 204–206 (2000).

445. Rege, T. A. & Hagood, J. S. Thy-1, a versatile modulator of signaling affecting cellular adhesion, proliferation, survival, and cytokine/growth factor responses. *Biochim. Biophys. Acta - Mol. Cell Res.* **1763,** 991–999 (2006).

446. Jiang, J. *et al.* Trastuzumab (herceptin) targets gastric cancer stem cells characterized by CD90 phenotype. *Oncogene* **31,** 671–682 (2012).

447. Kozii, R. *et al.* Thy-1 expression on blast cells from adult patients with acute myeloid leukemia. *Leuk. Res.* **21,** 381–5 (1997).

448. Ishiura, Y. *et al.* Anomalous expression of Thy1 (CD90) in B-cell lymphoma cells and proliferation inhibition by anti-Thy1 antibody treatment. *Biochem. Biophys. Res. Commun.* **396,** 329–334 (2010).

449. Takahashi, T. *et al.* Frequent expression of human Thy-1 antigen on pre-B cell acute lymphoblastic leukemia with t(9;22). *Int. J. Hematol.* **67,** 369–78 (1998).

450. Zhang, G. *et al.* Expression of Potential Cancer Stem Cell Marker ABCG2 is Associated with Malignant Behaviors of Hepatocellular Carcinoma. *Gastroenterol. Res. Pract.* **2013,** 782581 (2013).

451. Wee, B. *et al.* ABCG2 regulates self-renewal and stem cell marker expression but not tumorigenicity or radiation resistance of glioma cells. *Sci. Rep.* **6,** 25956 (2016).

452. Hu, J. *et al.* Expression of the cancer stem cell markers ABCG2 and OCT-4 in right-sided colon cancer predicts recurrence and poor outcomes. *Oncotarget* **8,** 28463–28470 (2017).

453. Sun, J. *et al.* A potentially functional polymorphism in ABCG2 predicts clinical outcome of non-small cell lung cancer in a Chinese population. *Pharmacogenomics J.* **17,** 280–285 (2017).

454. Mao, Q. & Unadkat, J. D. Role of the Breast Cancer Resistance Protein (BCRP/ABCG2) in Drug Transport—an Update. *AAPS J.* **17,** 65–82 (2015).

455. Kim, J. E. *et al.* Sonic hedgehog signaling proteins and ATP-binding cassette G2 are aberrantly expressed in diffuse large B-Cell lymphoma. *Mod. Pathol.* **22,** 1312–1320 (2009).

456. Turner, J. G. *et al.* ABCG2 expression, function, and promoter methylation in human multiple myeloma. *Blood* **108,** 3881–3889 (2006).

457. To, K. K. W., Zhan, Z. & Bates, S. E. Aberrant Promoter Methylation of the ABCG2 Gene in Renal Carcinoma. *Mol. Cell. Biol.* **26,** 8572–8585 (2006).

458. Boesch, M. *et al.* High prevalence of side population in human cancer cell lines. *Oncoscience* **3,** 85–87 (2016).

459. Hoe, S. L. L. *et al.* Evaluation of stem-like side population cells in a recurrent

nasopharyngeal carcinoma cell line. *Cancer Cell Int.* **14,** 101 (2014).

460. Mitsutake, N. *et al.* Characterization of Side Population in Thyroid Cancer Cell Lines: Cancer Stem-Like Cells Are Enriched Partly But Not Exclusively. *Endocrinology* **148,** 1797–1803 (2007).

461. Qi, W. *et al.* Sorting and identification of side population cells in the human cervical cancer cell line HeLa. *Cancer Cell Int.* **14,** 3 (2014).

462. Wu, C.-P. *et al.* Identification of Cancer Stem-Like Side Population Cells in Purified Primary Cultured Human Laryngeal Squamous Cell Carcinoma Epithelia. *PLoS One* **8,** e65750 (2013).

463. Jiang, Y., Gao, H., Liu, M. & Mao, Q. Sorting and biological characteristics analysis for side population cells in human primary hepatocellular carcinoma. *Am. J. Cancer Res.* **6,** 1890–1905 (2016).

464. Yoon, S.-O. *et al.* Down-regulation of CD9 expression and its correlation to tumor progression in B lymphomas. *Am. J. Pathol.* **177,** 377–86 (2010).

465. Yi, X.-J. *et al.* Aberrant Wnt/β-catenin signaling and elevated expression of stem cell proteins are associated with osteosarcoma side population cells of high tumorigenicity. *Mol. Med. Rep.* **12,** 5042–8 (2015).

466. Baron, B. W. *et al.* PIM1 gene cooperates with human BCL6 gene to promote the development of lymphomas. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 5735–9 (2012).

467. Kelly, G. L. *et al.* Targeting of MCL-1 kills MYC-driven mouse and human lymphomas even when they bear mutations in p53. *Genes Dev.* **28,** 58–70 (2014).

468. Giulino-Roth, L. *et al.* Targeted genomic sequencing of pediatric Burkitt lymphoma identifies recurrent alterations in antiapoptotic and chromatin-remodeling genes. *Blood* **120,** 5181–4 (2012).

469. Chen, Z. *et al.* LncFZD6 initiates Wnt/β-catenin and liver TIC self-renewal through BRG1-mediated FZD6 transcriptional activation. *Oncogene* **37,** 3098–3112 (2018).

470. Wagener, R. *et al.* Investigation of potential traces of pluripotency in germinal-center-derived B-cell lymphomas driven by MYC. *Blood Cancer J.* **5,** e317–e317 (2015).

471. Park, I. *et al.* Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* **423,** 302–305 (2003).

472. Molofsky, A. V. *et al.* Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. *Nature* **425,** 962–967 (2003).

473. Lessard, J. & Sauvageau, G. Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. *Nature* **423,** 255–260 (2003).

474. Kreso, A. *et al.* Self-renewal as a therapeutic target in human colorectal cancer. *Nat. Med.* **20,** 29–36 (2014).

475. WANG, Y. *et al.* Sorting, identification and enrichment of side population cells in THP-1 acute monocytic leukemia cells. *Oncol. Rep.* **29,** 1923–1931 (2013).

476. Zhou, J. *et al.* Side population rather than CD133+ cells distinguishes enriched tumorigenicity in hTERT-immortalized primary prostate cancer cells. *Mol. Cancer*

10, 112 (2011).

477. Salcido, C. D., Larochelle, A., Taylor, B. J., Dunbar, C. E. & Varticovski, L. Molecular characterisation of side population cells with cancer stem cell-like characteristics in small-cell lung cancer. *Br. J. Cancer* **102,** 1636–1644 (2010).

478. Xia, H. *et al.* Hepatocellular Carcinoma-propagating Cells are Detectable by Side Population Analysis and Possess an Expression Profile Reflective of a Primitive Origin. *Sci. Rep.* **6,** 34856 (2016).

479. Harrison, H. *et al.* Regulation of Breast Cancer Stem Cell Activity by Signaling through the Notch4 Receptor. *Cancer Res.* **70,** 709–718 (2010).

480. Holland, J. D., Klaus, A., Garratt, A. N. & Birchmeier, W. Wnt signaling in stem and cancer stem cells. *Curr. Opin. Cell Biol.* **25,** 254–264 (2013).

481. Hallett, R. M. *et al.* Small Molecule Antagonists of the Wnt/Beta-Catenin Signaling Pathway Target Breast Tumor-Initiating Cells in a Her2/Neu Mouse Model of Breast Cancer. *PLoS One* **7,** e33976 (2012).

482. Gargini, R., Cerliani, J. P., Escoll, M., Antón, I. M. & Wandosell, F. Cancer Stem Cell-Like Phenotype and Survival Are Coordinately Regulated by Akt/FoxO/Bim Pathway. *Stem Cells* **33,** 646–660 (2015).

483. Marhold, M. *et al.* HIF1α Regulates mTOR Signaling and Viability of Prostate Cancer Stem Cells. *Mol. Cancer Res.* **13,** 556–564 (2015).

484. Tavaluc, R. T., Hart, L. S., Dicker, D. T. & El-Deiry, W. S. Effects of Low Confluency, Serum Starvation and Hypoxia on the Side Population of Cancer Cell Lines. *Cell Cycle* **6,** 2554–2562 (2007).

485. Carmeliet, P. & Jain, R. K. Angiogenesis in cancer and other diseases. *Nature* **407,** 249–257 (2000).

486. Pouysségur, J., Dayan, F. & Mazure, N. M. Hypoxia signalling in cancer and approaches to enforce tumour regression. *Nature* **441,** 437–443 (2006).

487. Nazimek, K., Bryniarski, K., Santocki, M. & Ptak, W. Exosomes as mediators of intercellular communication: clinical implications. *Pol. Arch. Med. Wewn.* **125,** 370–80 (2015).

488. Hannafon, B. N. & Ding, W.-Q. Cancer stem cells and exosome signaling. *Stem cell Investig.* **2,** 11 (2015).

489. Murase, M. *et al.* Side population cells have the characteristics of cancer stem-like cells/cancer-initiating cells in bone sarcomas. *Br. J. Cancer* **101,** 1425–1432 (2009).

490. Hadnagy, A., Gaboury, L., Beaulieu, R. & Balicki, D. SP analysis may be used to identify cancer stem cell populations. *Exp. Cell Res.* **312,** 3701–3710 (2006).

491. Loebinger, M. R. *et al.* Squamous cell cancers contain a side population of stem-like cells that are made chemosensitive by ABC transporter blockade. *Br. J. Cancer* **98,** 380–387 (2008).

492. Moserle, L. *et al.* The Side Population of Ovarian Cancer Cells Is a Primary Target of IFN- Antitumor Effects. *Cancer Res.* **68,** 5658–5668 (2008).

493. Jakubikova, J. *et al.* Lenalidomide targets clonogenic side population in multiple myeloma: pathophysiologic and clinical implications. *Blood* **117,** 4409–19 (2011).

494. Huang, B. *et al.* Cancer Stem Cell-Like Side Population Cells in Clear Cell Renal Cell Carcinoma Cell Line 769P. *PLoS One* **8,** e68293 (2013).

495. Boesch, M. *et al.* The side population of ovarian cancer cells defines a heterogeneous compartment exhibiting stem cell characteristics. *Oncotarget* **5,** 7027–39 (2014).

496. Kim, K. H. *et al.* Side Population of Multiple Myeloma and Multiple Myeloma Stem Cell. *Blood* **124,** (2014).

497. Snijder, B. & Pelkmans, L. Origins of regulated cell-to-cell variability. *Nat. Rev. Mol. Cell Biol.* **12,** 119–125 (2011).

498. Lahm, H. *et al.* Autocrine growth factors secreted by the malignant human B-cell-line BJAB are distinct from other known cytokines. *Eur. Cytokine Netw.* **1,** 41–6

499. Grivennikov, S. I. IL-11: A Prominent Pro-Tumorigenic Member of the IL-6 Family. *Cancer Cell* **24,** 145–147 (2013).

500. Lehmann, B. D. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **121,** 2750–67 (2011).

501. Salem, M. L., El-Badawy, A. S. & Li, Z. Immunobiology and signaling pathways of cancer stem cells: implication for cancer therapy. *Cytotechnology* **67,** 749–759 (2015).

502. Xu, L. *et al.* Hypoxia-induced secretion of IL-10 from adipose-derived mesenchymal stem cell promotes growth and cancer stem cell properties of Burkitt lymphoma. *Tumor Biol.* **37,** 7835–7842 (2016).

503. Bolkun, L. *et al.* Involvement of BAFF and APRIL in Resistance to Apoptosis of Acute Myeloid Leukemia. *J. Cancer* **7,** 1979–1983 (2016).

504. Pelekanou, V. *et al.* BCMA (TNFRSF17) Induces APRIL and BAFF Mediated Breast Cancer Cell Stemness. *Front. Oncol.* **8,** 301 (2018).

505. Qin, H. *et al.* Novel BAFF-Receptor Antibody to Natively Folded Recombinant Protein Eliminates Drug-Resistant Human B-cell Malignancies *In Vivo*. *Clin. Cancer Res.* **24,** 1114–1123 (2018).

506. Cheson, B. D. Speed bumps on the road to a chemotherapy-free world for lymphoma patients. *Blood* **128,** 325–30 (2016).

507. Kolev, V. N. *et al.* PI3K/mTOR dual inhibitor VS-5584 preferentially targets cancer stem cells. *Cancer Res.* **75,** 446–55 (2015).

508. Frasson, C. *et al.* Inhibition of PI3K Signalling Selectively Affects Medulloblastoma Cancer Stem Cells. *Biomed Res. Int.* **2015,** 973912 (2015).

509. Zhu, M. M. *et al.* Increased JNK1 Signaling Pathway Is Responsible for ABCG2-Mediated Multidrug Resistance in Human Colon Cancer. *PLoS One* **7,** e41763 (2012).

510. Anoop, P. *et al.* Outcome of childhood relapsed or refractory mature B-cell non-

Hodgkin lymphoma and acute lymphoblastic leukemia. *Leuk. Lymphoma* **53,** 1882–1888 (2012).

511. Jourdain, A. *et al.* Outcome of and prognostic factors for relapse in children and adolescents with mature B-cell lymphoma and leukemia treated in three consecutive prospective "Lymphomes Malins B" protocols. A Société Française des Cancers de l'Enfant study. *Haematologica* **100,** 810–7 (2015).

512. Holohan, C., Van Schaeybroeck, S., Longley, D. B. & Johnston, P. G. Cancer drug resistance: an evolving paradigm. *Nat. Rev. Cancer* **13,** 714–726 (2013).

513. Ekhart, C., Rodenhuis, S., Smits, P. H. M., Beijnen, J. H. & Huitema, A. D. R. An overview of the relations between polymorphisms in drug metabolising enzymes and drug transporters and survival after cancer drug treatment. *Cancer Treat. Rev.* **35,** 18–31 (2009).

514. Fan, S. *et al.* p53 gene mutations are associated with decreased sensitivity of human lymphoma cells to DNA damaging agents. *Cancer Res.* **54,** 5824–30 (1994).

515. Fletcher, J. I., Haber, M., Henderson, M. J. & Norris, M. D. ABC transporters in cancer: more than just drug efflux pumps. *Nat. Rev. Cancer* **10,** 147–156 (2010).

516. Okosun, J. *et al.* Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat. Genet.* **46,** 176–181 (2014).

517. Jiang, Y. *et al.* Deep sequencing reveals clonal evolution patterns and mutation events associated with relapse in B-cell lymphomas. *Genome Biol.* **15,** 432 (2014).

518. Juskevicius, D., Dirnhofer, S. & Tzankov, A. Genetic background and evolution of relapses in aggressive B-cell lymphomas. *Haematologica* **102,** 1139–1149 (2017).

519. Aukema, S. M. *et al.* Sequential karyotyping in Burkitt lymphoma reveals a linear clonal evolution with increase in karyotype complexity and a high frequency of recurrent secondary aberrations. *Br. J. Haematol.* **170,** 814–825 (2015).

520. Kawashima, J., Okuno, H., Shiba, N., Yoshida, K. & Shiraishi, Y. Whole-Exome Sequencing in a Pediatric Patient with Relapsed/Refrac- tory Burkitt Leukemia Resistant to Conventional Chemotherapy. *Arch. Paediatr. Dev. Pathol.* **1**, 1001 (2017).

521. Sahlberg, S. H., Spiegelberg, D., Glimelius, B., Stenerlöw, B. & Nestor, M. Evaluation of Cancer Stem Cell Markers CD133, CD44, CD24: Association with AKT Isoforms and Radiation Resistance in Colon Cancer Cells. *PLoS One* **9,** e94621 (2014).

522. Bertolini, G. *et al.* Highly tumorigenic lung cancer CD133+ cells display stem-like features and are spared by cisplatin treatment. *Proc. Natl. Acad. Sci.* **106,** 16281–16286 (2009).

523. Dallas, N. A. *et al.* Chemoresistant Colorectal Cancer Cells, the Cancer Stem Cell Phenotype, and Increased Sensitivity to Insulin-like Growth Factor-I Receptor Inhibition. *Cancer Res.* **69,** 1951–1957 (2009).

524. Bao, S. *et al.* Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444,** 756–760 (2006).

525. Huang, B. *et al.* PKCε inhibits isolation and stemness of side population cells via the suppression of ABCB1 transporter and PI3K/Akt, MAPK/ERK signaling in renal cell carcinoma cell line 769P. *Cancer Lett.* **376,** 148–154 (2016).

526. Lee, K.-M. *et al.* MYC and MCL1 Cooperatively Promote Chemotherapy-Resistant Breast Cancer Stem Cells via Regulation of Mitochondrial Oxidative Phosphorylation. *Cell Metab.* **26,** 633–647 (2017).

527. Radpour, R. Tracing and targeting cancer stem cells: New venture for personalized molecular cancer therapy. *World J. Stem Cells* **9,** 169–178 (2017).

528. Newton, T. C., Wolcott, K. & Roberts, S. S. Comparison of the side populations in pretreatment and postrelapse neuroblastoma cell lines. *Transl. Oncol.* **3,** 246–51 (2010).

529. Fukunaga-Kalabis, M. & Herlyn, M. Beyond ABC: Another Mechanism of Drug Resistance in Melanoma Side Population. *J. Invest. Dermatol.* **132,** 2317–2319 (2012).

530. Hubbard, J. M. & Grothey, A. Napabucasin: An Update on the First-in-Class Cancer Stemness Inhibitor. *Drugs* **77,** 1091–1103 (2017).

531. *Intergroup trial for children or adolescents with B-Cell NHL or B-AL: evaluation of Rituximab efficacy and safety in high risk patients»*. NCT01516580 (2011).

532. Freireich, E. J., Gehan, E. A., Rall, D. P., Schmidt, L. H. & Skipper, H. E. Quantitative comparison of toxicity of anticancer agents in mouse, rat, hamster, dog, monkey, and man. *Cancer Chemother. reports* **50,** 219–44 (1966).

533. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27,** 1739–1740 (2011).

534. Gao, H. *et al.* High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nat. Med.* **21,** 1318–1325 (2015).

535. Tentler, J. J. *et al.* Patient-derived tumour xenografts as models for oncology drug development. *Nat. Rev. Clin. Oncol.* **9,** 338–350 (2012).

536. Rosfjord, E., Lucas, J., Li, G. & Gerber, H.-P. Advances in patient-derived tumor xenografts: From target identification to predicting clinical response rates in oncology. *Biochem. Pharmacol.* **91,** 135–143 (2014).

537. Hidalgo, M. *et al.* Patient-Derived Xenograft Models: An Emerging Platform for Translational Cancer Research. *Cancer Discov.* **4,** 998–1013 (2014).

538. Campbell, C. J. V. *et al.* The human stem cell hierarchy is defined by a functional dependence on Mcl-1 for self-renewal capacity. *Blood* **116,** 1433–1442 (2010).

539. Brasó-Maristany, F. *et al.* PIM1 kinase regulates cell death, tumor growth and chemotherapy response in triple-negative breast cancer. *Nat. Med.* **22,** 1303–1313 (2016).

540. Kuo, H.-P. *et al.* The role of PIM1 in the ibrutinib-resistant ABC subtype of diffuse large B-cell lymphoma. *Am. J. Cancer Res.* **6,** 2489–2501 (2016).

541. Gross, E. *et al.* B-chronic lymphocytic leukemia chemoresistance involves innate and acquired leukemic side population cells. *Leukemia* **24,** 1885–1892 (2010).

542. Milanovic, M. *et al.* Senescence-associated reprogramming promotes cancer stemness. *Nature* **553,** 96–100 (2017).

543. Dong, T. *et al.* The Expression of CD9 and PIK3CD is Associated with Prognosis of Follicular Lymphoma. *J. Cancer* **6,** 1222–1229 (2015).

544. Hemler, M. E. Tetraspanin functions and associated microdomains. *Nat. Rev. Mol. Cell Biol.* **6,** 801–811 (2005).

545. Arita, A. *et al.* Signaling pathways in lymphoma: pathogenesis and therapeutic targets. *Futur. Oncol.* **9,** 1549–1571 (2013).

546. Yoon, S.-O., Lee, I. Y., Zhang, X., Zapata, M. C. & Choi, Y. S. CD9 may contribute to the survival of human germinal center B cells by facilitating the interaction with follicular dendritic cells. *FEBS Open Bio* **4,** 370–6 (2014).

547. Maecker, H. T., Todd, S. C. & Levy, S. The tetraspanin superfamily: molecular facilitators. *FASEB J.* **11,** 428–42 (1997).

548. Zhang, C. *et al.* CD9 Blockade Effectively Suppresses Disease Progression of Pediatric B-Precursor Acute Lymphoblastic Leukemia and Enhances Chemosensitivity. *Blood* **130,** 1327 (2017).

549. Tripathi, V. *et al.* Long Noncoding RNA MALAT1 Controls Cell Cycle Progression by Regulating the Expression of Oncogenic Transcription Factor B-MYB. *PLoS Genet.* **9,** e1003368 (2013).

550. Yoshimoto, R., Mayeda, A. & Yoshida, M. MALAT1 long non-coding RNA in cancer. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1859,** 192–199 (2016).

551. Wu, X.-S. *et al.* MALAT1 promotes the proliferation and metastasis of gallbladder cancer cells by activating the ERK/MAPK pathway. *Cancer Biol. Ther.* **15,** 806–814 (2014).

552. Ying, L. *et al.* Upregulated MALAT-1 contributes to bladder cancer cell migration by inducing epithelial-to-mesenchymal transition. *Mol. Biosyst.* **8,** 2289 (2012).

553. Kim, S. H., Kim, S. H., Yang, W. I., Kim, S. J. & Yoon, S. O. Association of the long non-coding RNA MALAT1 with the polycomb repressive complex pathway in T and NK cell lymphoma. *Oncotarget* **8,** 31305–31317 (2017).

554. Wang, X. *et al.* LncRNA MALAT1 promotes development of mantle cell lymphoma by associating with EZH2. *J. Transl. Med.* **14,** 346 (2016).

555. Li, Q., Dai, Y., Wang, F. & Hou, S. Differentially expressed long non-coding RNAs and the prognostic potential in colorectal cancer. *Neoplasma* **63,** 977–983 (2016).

556. Zhang, H., Yang, F., Chen, S.-J., Che, J. & Zheng, J. Upregulation of long non-coding RNA MALAT1 correlates with tumor progression and poor prognosis in clear cell renal cell carcinoma. *Tumor Biol.* **36,** 2947–2955 (2015).

557. Schmidt, L. H. *et al.* The Long Noncoding MALAT-1 RNA Indicates a Poor Prognosis in Non-small Cell Lung Cancer and Induces Migration and Tumor Growth. *J. Thorac. Oncol.* **6,** 1984–1992 (2011).

558. Gutschner, T. *et al.* The Noncoding RNA MALAT1 Is a Critical Regulator of the Metastasis Phenotype of Lung Cancer Cells. *Cancer Res.* **73,** 1180–1189 (2013).

559. Latorre, E. *et al.* The Ribonucleic Complex HuR-MALAT1 Represses CD133 Expression and Suppresses Epithelial-Mesenchymal Transition in Breast Cancer. *Cancer Res.* **76,** 2626–2636 (2016).

560. Zeng, L., Cen, Y. & Chen, J. Long non-coding RNA MALAT-1 contributes to maintenance of stem cell-like phenotypes in breast cancer cells. *Oncol. Lett.* **15,** 2117–2122 (2017).

561. Ma, X.-Y. *et al.* Malat1 as an evolutionarily conserved lncRNA, plays a positive role in regulating proliferation and maintaining undifferentiated status of early-stage hematopoietic cells. *BMC Genomics* **16,** 676 (2015).

562. Sander, S. *et al.* MYC stimulates EZH2 expression by repression of its negative regulator miR-26a. *Blood* **112,** 4202–4212 (2008).

563. Schmitz, R. *et al.* Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N. Engl. J. Med.* **378,** 1396–1407 (2018).

564. Weiner, G. J. Rituximab: mechanism of action. *Semin. Hematol.* **47,** 115–23 (2010).

565. Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* **16,** 299–311 (2015).

566. Lo, A. & Qi, L. Genetic and epigenetic control of gene expression by CRISPR–Cas systems. *F1000Research* **6,** 747 (2017).

567. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159,** 647–661 (2014).

568. Yang, X. *et al.* A public genome-scale lentiviral expression library of human ORFs. *Nat. Methods* **8,** 659–661 (2011).

569. Perez-Pinera, P. *et al.* RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nat. Methods* **10,** 973–976 (2013).

570. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483,** 603–607 (2012).

571. Gentles, A. J. *et al.* The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat. Med.* **21,** 938–945 (2015).

572. Hummel, M. *et al.* A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N. Engl. J. Med.* **354,** 2419–30 (2006).

573. Ionov, Y. *et al.* Pim-1 protein kinase is nuclear in Burkitt's lymphoma: nuclear localization is necessary for its biologic effects. *Anticancer Res.* **23,** 167–78 (2003).

574. Wang, Z. *et al.* Pim-1: a serine/threonine kinase with a role in cell survival, proliferation, differentiation and tumorigenesis. *J. Vet. Sci.* **2,** 167–79 (2001).

575. Li, C.-M. *et al.* PEG10 is a c-MYC target gene in cancer cells. *Cancer Res.* **66,** 665–72 (2006).

576. García, J. L. *et al.* Abnormalities on 1q and 7q are associated with poor outcome in sporadic Burkitt's lymphoma. A cytogenetic and comparative genomic hybridization study. *Leukemia* **17,** 2016–2024 (2003).

577. DENG, X. *et al.* PEG10 plays a crucial role in human lung cancer proliferation,

progression, prognosis and metastasis. *Oncol. Rep.* **32,** 2159–2167 (2014).

578. LI, X. *et al.* PEG10 promotes human breast cancer cell proliferation, migration and invasion. *Int. J. Oncol.* **48,** 1933–1942 (2016).

579. Kainz, B. *et al.* Overexpression of the paternally expressed gene10 (PEG10) from the imprinted locus on chromosome 7q21 in high-risk B-cell chronic lymphocytic leukemia. *Int. J. Cancer* **121,** 1984–1993 (2007).

580. Xiong, J. *et al.* PEG10 promotes the migration of human Burkitt's lymphoma cells by up-regulating the expression of matrix metalloproteinase-2 and -9. *Clin. Invest. Med.* **35,** E117-25 (2012).

581. Sanulli, S. *et al.* Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation. *Mol. Cell* **57,** 769–783 (2015).

582. Cao, J., Li, H., Liu, G., Han, S. & Xu, P. Knockdown of JARID2 inhibits the proliferation and invasion of ovarian cancer through the PI3K/Akt signaling pathway. *Mol. Med. Rep.* **16,** 3600–3605 (2017).

583. Zhu, X.-X. *et al.* Jarid2 is essential for the maintenance of tumor initiating cells in bladder cancer. *Oncotarget* **8,** 24483–24490 (2017).

584. Rhyasen, G. W. & Starczynowski, D. T. IRAK signalling in cancer. *Br. J. Cancer* **112,** 232–237 (2015).

585. Loiarro, M., Ruggiero, V. & Sette, C. Targeting TLR/IL-1R signalling in human diseases. *Mediators Inflamm.* **2010,** 674363 (2010).

586. Ngo, V. N. *et al.* Oncogenically active MYD88 mutations in human lymphoma. *Nature* **470,** 115–119 (2011).

587. Miles, L. A., Garippa, R. J. & Poirier, J. T. Design, execution, and analysis of pooled *in vitro* CRISPR/Cas9 screens. *FEBS J.* **283,** 3170–3180 (2016).

588. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34,** 184–191 (2016).

589. Piccaluga, P. P. *et al.* Gene expression analysis uncovers similarity and differences among Burkitt lymphoma subtypes. *Blood* **117,** 3596–608 (2011).

590. Guikema, J. E. J., Schuuring, E. & Kluin, P. M. Structure and Consequences of IGH Switch Breakpoints in Burkitt Lymphoma. *JNCI Monogr.* **2008,** 32–36 (2008).

591. Cobaleda, C., Jochum, W. & Busslinger, M. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature* **449,** 473–477 (2007).

592. Palmeira, A., Sousa, E., Vasconcelos, M. H. & Pinto, M. M. Three decades of P-gp inhibitors: skimming through several generations and scaffolds. *Curr. Med. Chem.* **19,** 1946–2025 (2012).

593. Li, W. *et al.* Overcoming ABC transporter-mediated multidrug resistance: Molecular mechanisms and novel therapeutic drug strategies. *Drug Resist. Updat.* **27,** 14–29 (2016).

594. Jaramillo, A. C., Al Saig, F., Cloos, J., Jansen, G. & Peters, G. J. How to overcome ATP-binding cassette drug efflux transporter-mediated drug resistance? *Cancer*

Drug Resist **1,** 6–29 (2018).

Appendix 1:

CRISPR Cas9 Screen Hits

Chemotherapy	Rituximab	Chemotherapy + Rituximab
ABCB1	ADH1C	ABCB1
ABCC1	AGAP3	ABCC1
ACTR2	AKT1	ACTR2
ADAM9	AKT2	ADAM9
AKT1	ALCAM	AKT1
ALCAM	ASAP3	ALCAM
ARGEF1	ATAD3A	ARGEF1
ARHGAP21	ATG10	ARHGAP21
ARHGAP21	ATP5S	ASAP3
ARHGDIA	ATXN7L2	BACE2
ASAP3	C1QL4	BCL2
BACE2	CACHD1	BRCC3
BCL2	CACNA2D3	CADM2
BRCC3	CACNG5	CALCR
C6orf1	CADM2	CASP6
CADM2	CAPNS2	CBX3
CALCR	CELA2A	CD3G
CASP6	CITED2	CDC25B
CBX3	CMTM6	CDK4
CDC20B	COLEC11	DDB2
CDK4	CXCL11	DDX55
DDB2	CYP46A1	DKK3
DDX55	DCAF16	EPB41
DIO3	DCUN1D4	EPHA1
DKK3	DEFB110	FABP4
DLX6	FAM110A	FAM86B2
EGFR	FBXO36	FRMD3
EPHA1	FXR2	GLB1
FABP4	GNA12	GNA14
FAM24B	GPR65	GRB2
FAM86B2	HAUS1	HR
FBXO27	HAUS8	IKZF2
FRMD3	HAX1	ITGB1
FZD6	HPN	JARID2
GFI1	HTN3	KRAS
GLB1	HTR3D	LIN7B
GNA14	IL10RA	MAD2L2

GRB2	IL36B	MALT1
HDAC6	IRAK1	MAPK1
HDAC9	ITGB1	MCL1
HR	JPH1	MMRN2
IKZF2	KCNG3	NFATC4
ITGA6	KHDRBS1	NFKB1
ITGB1	KIAA0930	OR10G7
JARID2	KRAS	PARD6B
KRAS	KRT3	PDCD1
LIN7B	KRT6B	PDS5B
MAD2L2	LHFPL3	PEG10
MCL1	LIMK1	PIM1
MMRN2	LIN7B	PITX1
NFATC4	LMCD1	RAC1
NFKB1	LMX1A	RBM24
OR10G7	MAD2L2	RLIM
OR6C76	MAP3K15	STAT5B
PARD6B	MAPK1	TTC5
PDE6H	MAPK13	
PDS5B	MPZL1	
PEG10	MTX2	
PIM1	MZT2B	
PITX1	NADK2	
RAC1	NFKB1	
RAF1	NIPAL4	
RBM24	NUP210	
RIC8B	NXT2	
RLIM	PCDH1	
RLIM	PHF21B	
SNCB	PRDM16	
STAT5B	PRDM2	
STAT6	PRPF40B	
TTC5	RAB12	
WNT5B	RAC1	
ZNF254	RAC2	
	RGS5	
	RIBC2	
	ROM1	
	RPS6KC1	
	S1PR5	
	SAMD1	

	SEMA3F	
	SH3BGR	
	SLC14A2	
	SLC29A1	
	SLITRK4	
	SOX14	
	TAF11	
	TMEM106C	
	TMEM141	
	TMEM204	
	TMX1	
	TP53BP2	
	USP10	
	XRCC6	